THE PROCEEDINGS OF
THE PHYSICAL SOCIETY
Section B

VOL. 63, PART 2 1 February 1950 No. 362 B

The Properties of Ferromagnetic Compounds at Centimetre Wavelengths

BY J. B. BIRKS
Department of Natural Philosophy, University of Glasgow

MS. received 29th June 1949

ABSTRACT. The magnetic and dielectric properties of γ-ferric oxide, magnetite, Mn–Zn ferrite and Ni–Zn ferrite have been measured at wavelengths from 60 cm. to 1·23 cm. The pronounced magnetic dispersion and absorption which occur are attributed to the decay of the rotational magnetization, due to Larmor spin resonance in the internal anisotropy field H_N. The theoretical relation $H_N=2K/M$, where K is the anisotropy coefficient, and M the saturation intensity, agrees satisfactorily with the observations, and $|K|$ has been derived for each material. An inverse relationship between H_N and the initial susceptibility is shown to account quantitatively for the magnetic resonance observed in high-permeability materials at much lower frequencies.

§ 1. INTRODUCTION

An investigation has been made of the magnetic and dielectric properties of four ferromagnetic compounds at wavelengths from 60 cm. to 1·23 cm., using the waveguide dual-impedance method, previously described (Birks 1948 a). The complex permeability and permittivity of the materials are obtained from observations of the input impedances of thin specimens, fitted into coaxial or rectangular waveguides, when terminated in a short- and an open-circuit. The ferromagnetic compounds are of interest because, due to their low conductivity, the high-frequency field is able to penetrate much more deeply than in a metal. It is thus possible to observe magnetic dispersion and absorption effects due to internal phenomena, which are obscured in metals by the intense skin-effect.

Almost all the previous experimental work on high-frequency permeability has been concerned with metals (Allanson 1945) and several attempts have been made to interpret the observed dispersion in terms of various internal relaxation and resonance mechanisms (Arkadiew 1945), analogous to those occurring in polar dielectric materials (Debye 1929) and in paramagnetic salts (Gorter 1947). The major magnetic dispersion in ferromagnetic metals, which occurs at frequencies above 100 Mc/s., is not, however, explicable in such terms, as the skin-effect restricts the penetration of the oscillatory field to a thin surface layer of less than a single domain thickness. Kittel (1946) has proposed an alternative magnetic dispersion mechanism for metals, based on this incomplete penetration of surface domains. Measurements on the microwave properties of carbonyl iron powder give results in agreement with Kittel's theory (Birks 1948 b).

Since the microwave magnetization of metals is only a surface effect, it appears that if high-frequency measurements are to be used to study the internal dynamics of ferromagnetic materials, they must be made on materials of conductivity sufficiently low for the skin-depth to be large compared with the domain size. The ferromagnetic compounds—γ-ferric oxide, magnetite and the ferrites—satisfy this criterion. The magnetic dispersion and absorption observed in these materials can hence only arise from internal restraints on the magnetization (Birks 1948 a).

§ 2. EXPERIMENTAL RESULTS

(i) *Typical γ-ferric Oxide Mixture*

The ferromagnetic compounds, in the form of finely-divided powders, were mixed with paraffin wax to form solid samples for measurement. A typical series of observations on a single mixture, containing 32·6% by volume of γ-ferric oxide in wax, will be described initially to illustrate the general nature of the results. The complex permeability $\mu = \mu' - i\mu''$ and the complex permittivity $\epsilon = \epsilon' - i\epsilon''$ of the mixture were measured at 9 different wavelengths: 58·5, 39·2, 29·8, 22·4, 15·3, 8·93, 5·97, 3·09 and 1·23 cm. Measurements were made at each wavelength on three or more specimens, moulded from the same material. The mean values of the components of μ and ϵ are plotted in Figures 1 and 2. The initial static permeability μ_S of the mixture, measured by a ballistic method, was 3·2.

Although the magnetic properties are of primary interest, the simultaneous derivation of ϵ from the same set of observational data provides a check on the validity of the measurements. The gradual decrease of permittivity observed is typical of the dielectric dispersion of solid inorganic compounds. Unusual features in the magnetic dispersion cannot, therefore, be attributed to experimental error. This point is emphasized, because certain abnormal magnetic dispersion phenomena reported in metals (Kartschagin 1922, Gans and Loyarte 1921) were later shown to arise from errors of observation (Wait 1927).

The magnetic dispersion and absorption (Figure 1) is pronounced, and displays several interesting features. On the long-wavelength side of the dispersion region ($\lambda \sim 150$ cm.), the limiting value of μ', $\mu_0 = 3{\cdot}05$ is within 5% of μ_S, the static permeability, and hence no appreciable dispersion can occur in the interval from $\lambda = \infty$ to $\lambda \sim 150$ cm. At this latter wavelength μ' begins to decrease, becoming unity at $\lambda \sim 7{\cdot}5$ cm., while $\tan \delta_\mu (= \mu''/\mu')$ rises to a maximum at about the same wavelength. At shorter wavelengths, $\tan \delta_\mu$ decreases again rapidly, and μ' falls below unity, corresponding to a negative susceptibility. At $\lambda = 1{\cdot}23$ cm. the magnetic loss is negligible, while μ' is approximately unity, which is the ultimate infra-red value for ferromagnetic materials (Hagen and Rubens 1903). Thus, practically all the magnetic dispersion and absorption occurs in the wavelength interval 150 cm. to 1 cm.

Further measurements were made to determine the effect of a static magnetic field on the observed dispersion. Thin discs (thickness/diameter $\simeq 0{\cdot}1$) of the materials were fitted normally in the coaxial line impedance apparatus, and the static magnetic field H was applied parallel to the faces of the discs. Specimens of similar dimensional ratios were used to avoid differences in demagnetization coefficients. The effective values of μ as a function of H, up to 2,500 gauss, were measured at wavelengths from 58·5 cm. to 5·97 cm. The permittivity ϵ was unaffected by H. The static magnetization of the mixture was also measured by a ballistic method, and the incremental static permeability $\mu_S (= dB/dH)$ and its reversible component μ_{SR} were derived as functions of the applied field H.

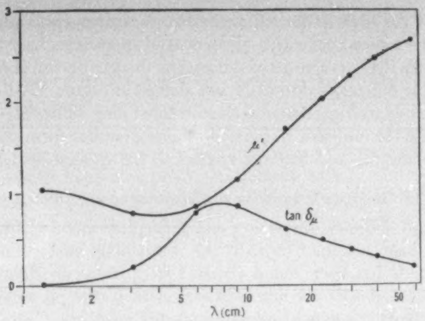


Figure 1. Magnetic dispersion and absorption of 32·6% γ-ferric oxide mixture.

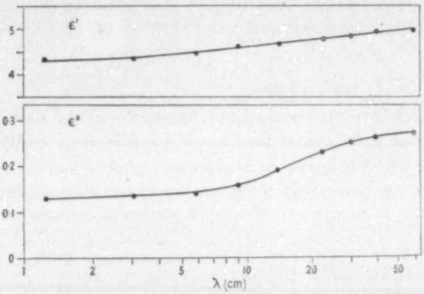


Figure 2. Dielectric dispersion and absorption of 32·6% γ-ferric oxide mixture.

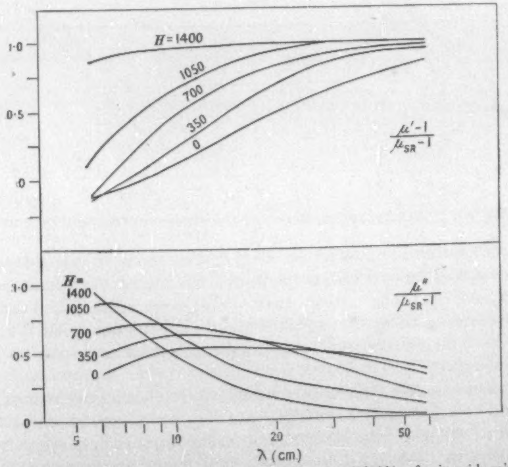


Figure 3. Effect of applied field H on magnetic dispersion (32·6% γ-ferric oxide mixture).

The results of the high-frequency and static measurements, adjusted to allow for the demagnetization coefficients of the coaxial specimens, have been published previously (Birks 1947). The derived magnetic dispersion and absorption curves, corresponding to different values of H, are plotted in Figure 3. For comparison, the ordinates are expressed relative to the corresponding values of $\mu_{SR}-1$. It will be observed that the magnetic dispersion retains a similar form, but that it shifts to shorter wavelengths, and becomes steeper, as the applied field H is increased.

(ii) *Magnetic Dispersion of Ferromagnetic Compounds*

Magnetic and dielectric dispersion and absorption curves, similar to Figures 1 and 2, have been obtained for mixtures containing other concentrations of γ-ferric oxide. It has been found (Birks 1948 a) that, at each wavelength of measurement, μ and ϵ of the mixtures vary with v, the proportion by volume of the ferromagnetic compound, according to the relations

$$\log|\mu| = v\log|\mu_a|, \qquad \ldots\ldots(1)$$

$$\tan\delta_\mu = v\tan\delta_{\mu_a}, \qquad \ldots\ldots(2)$$

$$\log|\epsilon| = v\log|\epsilon_a| + (1-v)\log|\epsilon_w|, \qquad \ldots\ldots(3)$$

$$\tan\delta_\epsilon = v\tan\delta_{\epsilon_a}, \qquad \ldots\ldots(4)$$

where suffix a refers to the (extrapolated) properties of the ferromagnetic solid, and suffix w to the properties of the wax ($|\epsilon_w| = 2{\cdot}28$, $|\mu_w| = 1{\cdot}00$, $\delta_{\mu_w} = \delta_{\epsilon_w} = 0$.) These relations have been used to extrapolate the observations on the mixtures to zero dilution, and thus obtain the properties of the ferromagnetic solid. The derived magnetic dispersion and absorption curves of γ-ferric oxide are plotted in Figure 4.

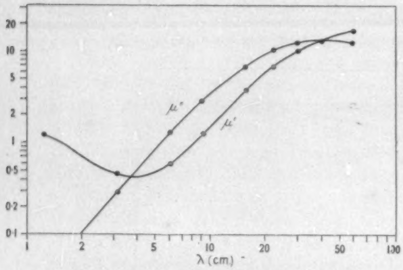


Figure 4. Magnetic spectrum of γ-ferric oxide.

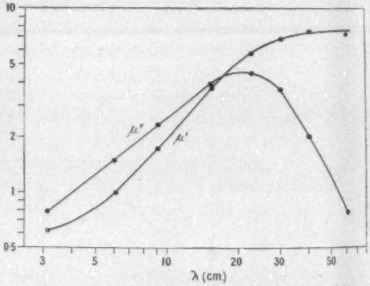


Figure 5. Magnetic spectrum of magnetite.

Observations have been made on three similar series of mixtures containing various concentrations of magnetite, nickel-zinc ferrite and manganese-zinc ferrite respectively. The mixture laws (1)–(4) were again found to be valid. The derived magnetic spectra of these three materials are plotted in Figures 5, 6 and 7. The same type of broad magnetic resonance dispersion occurs in each compound.

The values of μ_0, the limiting permeability at frequencies below the dispersion region, and of μ_S, the initial static permeability, have been obtained from the dispersion curves, and from static measurements on mixtures respectively, and they are listed in Table 1.

§ 3. THEORY OF MAGNETIC DISPERSION

(i) *Nature of Magnetization Process*

The domain theory of ferromagnetism (Becker and Döring 1939) differentiates four elementary processes which occur in the static magnetization of a demagnetized material.

I. Reversible displacements of the Bloch walls separating adjacent domains.
II. Irreversible displacements of the domain walls.
III. Irreversible rotations of the spins within a domain from one direction of easy magnetization to another.
IV. Reversible rotations of the domain spins towards the direction of the applied field.

The reversible translational magnetization I takes place in weak fields only, and is generally considered responsible for the initial permeability. The irreversible processes II and III, responsible for the Barkhausen effect, occur in medium fields over the steep portion of the *B–H* curve. The reversible rotational magnetization IV can occur at all field strengths, and is the only process operative in high fields.

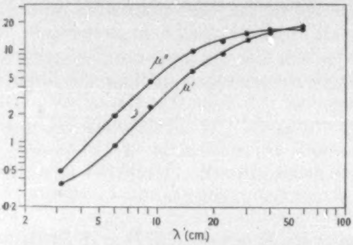


Figure 6. Magnetic spectrum of nickel-zinc ferrite.

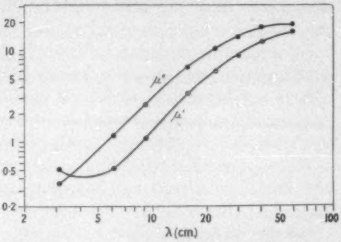


Figure 7. Magnetic spectrum of manganese-zinc ferrite.

Table 1

Material	μ_S	μ_0	λ_N (cm.)	ν_N (Mc/s.)	H_N (gauss)	M	$\lvert K\rvert$ (ergs/cm^3)
Magnetite	15	8	18	1670	590	450	$1{\cdot}3\times10^5$
γ-ferric oxide	34	30	40	750	270	350	$4{\cdot}7\times10^4$
Ni–Zn ferrite	40	34	50	600	210	300	$3{\cdot}2\times10^4$
Mn–Zn ferrite	48	44	70	430	150	200	$1{\cdot}5\times10^4$

Becker (1938) has developed a theory of magnetic dispersion for the translational and irreversible magnetization components. The dispersion is attributed to the damping action of eddy current fields, induced in the vicinity of domains by the extension of their boundaries or by irreversible rotations. The disappearance of the irreversible magnetization components at relatively low frequencies, predicted by Becker, has been confirmed experimentally (Birks 1947). Hence, only reversible processes are operative in the microwave region. For the relaxation of the reversible translational magnetization, Becker obtains a Debye-type equation of the form

$$\frac{\mu-1}{\mu_S-1}=\frac{1}{1-i\nu/\nu'} \qquad \ldots\ldots(5)$$

for the permeability μ in weak oscillatory fields of frequency ν, where ν' is the relaxation frequency. The form of (5) is identical with the dispersion relation obtained by Arkadiew (1913) from his semi-empirical theory of magnetic viscosity. The observed dispersion and absorption curves differ appreciably from (5). The (μ', λ) curves are steeper, the (μ'', λ) curves are sharper, and $\mu'-1$ becomes negative, while (5) gives positive values only. The permeability does not, therefore, decay by a relaxation process, as would be expected if it were due to translational magnetization.

The observations on the effect of an applied field (Figure 3) show that a similar type of dispersion occurs when high static fields are present. Process I, which is restricted to weak fields, cannot be responsible, and hence the magnetic dispersion must be accounted for in terms of the one alternative process, reversible rotational magnetization, which is operative at all field strengths.

Comparison of μ_0 and μ_S (Table 1) shows that, with the exception of magnetite, practically all the dispersion occurs in the microwave region. Since this dispersion is attributable to rotational magnetization, it follows that the translational magnetization in these materials, even in static fields, is relatively small. This is probably due to their low conductivity, which will inhibit the spread of eddy-currents associated with domain-boundary displacements. Magnetite, in which the initial rotational and translational components are of equal magnitudes ($\mu_S \simeq 2\mu_0$), has a higher conductivity than the other three compounds. The magnetic dispersion observed in magnetite by Zimowski (1937) at 5–20 Mc/s. probably corresponds to the decay of the translational magnetization.

(ii) *Natural Ferromagnetic Resonance*

The rotation of the spins within a domain towards the direction of the applied field is opposed by the crystalline anistropy forces, which tend to keep the spins aligned along an easy magnetization axis, O*n*. Landau and Lifshitz (1935), in considering a theoretical domain model of a ferromagnetic crystal, have represented these anisotropy forces by an equivalent internal magnetic field H_n, acting along O*n*. This natural anisotropy field, which is of the order of 1,000 gauss, arises from spin–orbit interaction, and is distinct from the much larger Weiss molecular field, arising from the spin exchange forces. Landau and Lifshitz have considered the effect of a weak oscillatory magnetic field applied transverse to O*n*, so that spin rotation is the only magnetization process possible. They find that magnetic dispersion takes place as a resonance process, resonance occurring at the Larmor precession frequency ν_n of the electron spins in the internal field, given by

$$h\nu_n=g\mu_B H_n, \qquad \ldots\ldots(6)$$

where g is the Landé factor (= 2 for electron spins), and μ_B is the Bohr magneton.

The differences, noted above, between the observed dispersion and the relaxation equation (5), and, in particular, the negative values of $\mu'-1$, are characteristic of a resonance effect. These observations provide direct experimental evidence of the existence of an internal field. Any magnetic resonance process occurring in the microwave region must be associated with a field of the order of 1,000 gauss, since for electron spins (6) becomes numerically

$$H_n\lambda_n \doteq 10{\cdot}7 \times 10^3 \text{ gauss cm.}, \qquad \ldots\ldots(7)$$

where $\lambda_n (= \mathbf{c}/\nu_n)$ is the resonant wavelength. Since magnetic resonance is observed when no such field is applied externally, an effective field H_n must exist within the ferromagnetic material. When an external field H is applied to the material, it will increase the total field acting on the spins, and the resonance will shift to higher frequencies, as observed (Figure 3). When $H > H_n$, the resonant frequency will be determined primarily by the external field, and induced ferromagnetic resonance will occur (Birks 1948 c). Thus, qualitatively, the observations are consistent with Landau and Lifshitz's hypothesis of the existence of an internal field. The identification of this field with that arising from anisotropy will be discussed later.

(iii) *Natural Resonance in Polycrystalline Materials*

The conditions within each domain or region of a polycrystalline material, in which a unidirectional internal field H_n exists, will be similar to those in a paramagnetic or saturated ferromagnetic material in which the spins are aligned by an external polarizing field H_n. In this latter case, when an oscillatory field is applied transverse to H_n, induced magnetic resonance occurs at the Larmor frequency ν_n of the spins in the external field (Zavoisky 1946). This effect is thus closely related to the natural magnetic resonance, due to the internal field. The resonance dispersion equation, derived by Frenkel (1945),

$$\frac{\mu - 1}{\mu_0 - 1} = \frac{\nu_n^2}{\nu_n^2 - \nu^2 + 2i\nu\nu'}, \qquad \ldots\ldots(8)$$

where ν' is the damping frequency, μ_0 the initial rotational permeability, and μ the permeability at frequency ν, has been verified experimentally for a saturated ferromagnetic by Yager and Bozorth (1947). The experiments on induced ferromagnetic resonance show that $\nu' \ll \nu_n$, and, hence, from (8), the maximum absorption μ'' occurs when $\nu = \nu_n$.

In an unmagnetized polycrystalline material the internal field H_n will not be uniform, but will vary in direction, due to local changes in domain and crystal structure, and in magnitude, due to magnetic and thermal interactions. These will lead to a broad distribution of values of H_n about a mean H_N, corresponding to the undisturbed, unicrystalline field. The natural polycrystalline magnetic dispersion thus consists of a broad spectrum of relatively sharp resonances ($\nu' \ll \nu_n$) occurring locally, and distributed about a mean resonant frequency ν_N. The over-simplification of applying (8) directly to the dispersion (Birks 1948 c)* is invalid, since it involves postulating values of $\nu' > \nu_n$ to account for the breadth of the resonance. In the absence of a suitable theory for treating the magnetic interactions, no attempt has yet been made to extend (8) to the polycrystalline case, though a tentative approach to the problem has been made recently by

---

* The larger values of H_n obtained previously, on the assumption, from (8), that the resonant wavelength is that at which $\mu' = 1$, were in error by a factor of ~ 10. The corrected anisotropy fields are insufficient to account for more than a small fraction of the g-value anomaly ($g > 2{\cdot}0$) observed in induced ferromagnetic resonance (Yager and Bozorth 1947). This anomaly is a distinct effect, and is not, as previously suggested (Birks 1948 c), due merely to anisotropy.

Crouch (1949). The mean resonant wavelength λ_N can be derived directly from the observed peak of magnetic absorption. The values of λ_N, ν_N and H_N obtained are given in Table 1.

The observations on the effect of an applied field (Figure 3) agree qualitatively with the theory proposed. Apart from the shift towards higher frequencies due to the increase of the effective field, there is a reduction in the breadth of the resonance, corresponding to the decrease in the magnetic disorder in the material as the domain moments are turned irreversibly from random orientations towards the direction of the applied field.

§4. THE INTERNAL FIELD

(i) *Anisotropy Energy and the Internal Field*

The anisotropic properties of a ferromagnetic material are normally described in terms of an anisotropy energy coefficient K. For a cubic crystal, magnetized to saturation, the component of the magnetic energy due to anisotropy may be written

$$f = K(\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2), \qquad \ldots\ldots(9)$$

where α_1, α_2, α_3 are the direction cosines of the magnetization relative to the (100) axes.

The internal anisotropy field H_N is related to K and the saturation intensity M. Let us consider a single domain of a cubic crystal, lying in the [100] plane $0yz$, whose easy direction of magnetization lies along a (100) axis, $0z$ (Figure 8). The internal field H_N acts along $0z$, and in the unmagnetized state M is in the same direction. If a small transverse field H_x is applied along $0x$, M will rotate through an angle θ in the plane $0xz$. The internal field H_N exerts a restoring torque on M, given by

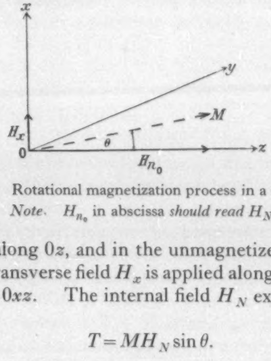


Figure 8. Rotational magnetization process in a single domain.
Note. H_{n_0} in abscissa *should read* H_N.

$$T = MH_N \sin\theta. \qquad \ldots\ldots(10)$$

Alternatively, if the process is described in terms of f, in the plane $0xz$,

$$f = K\cos^2\theta\sin^2\theta \qquad \ldots\ldots(11)$$

and the anisotropy energy exerts a restoring torque

$$T = \frac{\partial f}{\partial\theta} = K\sin 2\theta\cos 2\theta. \qquad \ldots\ldots(12)$$

Equating (10) and (12), for θ small, we obtain

$$H_N = \frac{2K}{M}. \qquad \ldots\ldots(13)$$

For a crystal in which the (111) axes are preferred, a similar analysis gives

$$H_N = -\frac{2K}{M}. \qquad \ldots\ldots(14)$$

The value of K is normally derived experimentally from the differences between the energies of magnetization of a single crystal along its different axes. Quittner (1909) has measured the magnetization curves of a magnetite crystal, and a value of $K = -1{\cdot}124 \times 10^5$ ergs/cm³ has been derived from his data. Magnetite, like the other compounds studied, has a spinel-type cubic structure. Its preferred direction of magnetization is along the cube diagonal. The saturation intensity of artificial magnetite at room temperature, $M = 450$ (Weiss and Forrer 1929) and hence from (14) the theoretical value of the anisotropy field $H_N = 500$. This agrees satisfactorily with that for the internal field, $H_N = 590$, obtained from the natural magnetic absorption peak.* It is concluded therefore that the internal field responsible for the magnetic resonance does arise from the ferromagnetic anisotropy.

No data are available on the magnetic properties of single crystals of the other compounds studied. The magnitude of the anisotropy constant $|K|$ can be estimated from the experimental values of H_N and M, using (13) or (14). The sign of K will depend on the preferred directions of magnetization. The values of $|K|$ obtained are listed in Table 1. The magnitude of K is one of the principal factors influencing the initial permeability, a small $|K|$ being associated with a large μ_S. It will be observed from Table 1 that the two quantities are found to be in inverse sequence.

(ii) *The Initial Permeability and the Internal Field*

Comparison of the initial permeabilities of the four compounds with those from other sources (Welo and Baudisch 1925, Kittel 1946, Snoek 1947) indicates that, with the exception of the manganese-zinc ferrite, the specimens are representative. The properties of the Mn–Zn ferrite specimen, however, differ appreciably from the optimum mixtures obtained by Snoek (1947, 1948), which have $\mu_S \simeq \mu_0 \simeq 1{,}000$. In such materials $|K|$ will be much smaller, and, consequently, natural magnetic resonance should occur at much lower frequencies. Snoek (1948) has observed the initial part of a sharp rise in the magnetic absorption of optimum Mn–Zn ferrites at frequencies below 1 Mc/s. This is interpreted by Snoek as the fringe of a magnetic resonance region, but his published observations are confined to frequencies well below the natural resonance frequency. From the limited data it is estimated that this is $\simeq 20$ Mc/s., which is equivalent, on the theory given above, to $H_N \simeq 7$ gauss and $|K| \simeq 700$ ergs/cm³.

A simple model, similar to that of Figure 8, has been used by Snoek (1948) to relate the initial rotational susceptibility X_0 and the internal field H_N. Neglecting all interactions, and assuming that H_N is constant throughout the polycrystalline material, it is found that

$$H_N = \frac{2M}{3X_0} = \frac{8\pi M}{3(\mu_0 - 1)}, \qquad \ldots\ldots(15)$$

---

* The residual discrepancy between theory and experiment is reduced, if a value of g greater than 2·0, as indicated by induced ferromagnetic resonance experiments (Yager and Bozorth 1947, Birks 1948 c) is adopted ; e.g. $g = 2{\cdot}2$, $H_N = 536$; $g = 2{\cdot}36$, $H_N = 500$.

so that for a given value of M, H_N should be inversely proportional to $\mu_0 - 1$. The parameters derived for the ferrite, $\mu_0 \simeq 1{,}000$, can be compared with those for the specimen, $\mu_0 \simeq 44$. The values of H_N are approximately in the ratio 1 to 22, while those of $\mu_0 - 1$ are in the ratio 23 to 1. Thus the relative magnitudes of the internal fields and the resonant frequencies are simply accounted for by the difference in the initial permeabilities of the two specimens.

In Table 2 the theoretical values of H_N from M, μ_0 and (15) are compared with the experimental results. There is close agreement in the case of magnetite, but for the other materials there is a discrepancy by a factor of the order of 3. This order of agreement is satisfactory, however, considering the extreme simplicity of the theoretical model and the neglect of all interactions in the derivation of (15).

Table 2

Material	H_N Expt.	H_N Theory (15)	$\frac{H_N\text{ (exp)}}{H_N\text{ (theor)}}$	μ_0
Magnetite	590	540	1·1	8
γ-ferric oxide	270	100	2·7	30
Ni–Zn ferrite	210	74	2·9	34
Mn–Zn ferrite	150	42	3·6	44

ACKNOWLEDGMENTS

The author wishes to acknowledge the tenure of an I.C.I. Research Fellowship during the initial stages of this work, and also to thank Professor P. I. Dee for the use of the facilities of his laboratory.

REFERENCES

Allanson, J. T., 1945, *J. Instn. Elect. Engrs.*, **92,** III, 247.

Arkadiew, W., 1913, *Phys. Z.*, **14,** 928 ; 1945, *J. Phys. U.S.S.R.*, **9,** 373.

Becker, R., 1938, *Phys. Z.*, **39,** 856.

Becker, R., and Döring, W., 1939, *Ferromagnetismus* (Berlin : Springer).

Birks, J. B., 1947, *Nature, Lond.*, **160,** 535 ; 1948 a, *Proc. Phys. Soc.*, **60,** 282 ; 1948 b, *Phys. Rev.*, **74,** 843 ; 1948 c, *Ibid.*, **74,** 988.

Crouch, G. E., 1949, *Phys. Rev.*, **75,** 525.

Debye, P., 1929, *Polar Molecules* (New York : Chemical Catalog Co.).

Frenkel, J., 1945, *J. Phys. U.S.S.R.*, **9,** 229.

Gans, R., and Loyarte, R. G., 1921, *Ann. Phys., Lpz.*, **64,** 209.

Gorter, C. J., 1947, *Paramagnetic Relaxation* (Amsterdam : Elsevier).

Hagen, E., and Rubens, H., 1903, *Ann. Phys., Lpz.*, 11, 873.

Kartschagin, K., 1922, *Ann. Phys., Lpz.*, **67,** 325.

Kittel, C., 1946, *Phys. Rev.*, **70,** 281.

Landau, L., and Lifshitz, E., 1935, *Phys. Z. Sowjet*, **8,** 153.

Quittner, V., 1909, *Ann. Phys., Lpz.*, **30,** 289.

Snoek, J. L., 1947, *New Developments in Ferromagnetic Materials* (Amsterdam : Elsevier) ; 1948, *Physica*, **14,** 207.

Wait, G. R., 1927, *Phys. Rev.*, **29,** 566.

Weiss, P., and Forrer, R., 1929, *Ann. Phys., Paris*, **12,** 279.

Welo, L. A., and Baudisch, O., 1925, *Phil. Mag.*, **50,** 399.

Yager, W. A., and Bozorth, R. M., 1947, *Phys. Rev.*, **72,** 80.

Zavoisky, E., 1946, *J. Phys. U.S.S.R.*, **10,** 197.

Zimowski, J., 1937, *Acta phys. polon.*, **6,** 6.

The Field in an Electron–Optical Immersion Objective

By L. JACOB
I.C.I. Fellow, The University, Manchester *

Communicated by J. Rotblat ; *MS. received* 4*th August* 1948,
in amended form 7*th July* 1949

ABSTRACT. The electrolytic tank method was developed to record rapidly the potential variations which occur in the field of the immersion objective in the absence of space charge. The axial distribution over the region explored was found to conform with a law of the form $V = Ae^{kz}$. Here A is a voltage scale factor while k is a geometrical scale factor. The experimental results appear to show that for constant black-out voltage the gradient at the cathode remains constant over part of the range explored, independent of geometry. The ratio of accelerator to modulator voltage for zero gradient at the cathode varies inversely as the square of the hole diameter. A determination of the emitting area of the cathode was also made from the field plots.

---

§ 1. INTRODUCTION

The electron optical immersion objective derives its name from the immersion objective of 'light' optics in that the electron-emitting area is immersed in, and forms part of, the lens field. In a wide sense it forms an integral part of every electron-emitting device in which electrons are drawn away to one or more electrodes by an electric field. Some of its focal properties, including aberration effects, have been described by Johannson (1933), and a study of its action as a gun for cathode-ray tubes can be found in the book by Maloff and Epstein (1938). Recently, very useful empirical design curves have been published by Moss (1945) which enable beam characteristics to be controlled.

The fundamental problems involved are those associated with the behaviour of electrostatic fields towards electron beams entering them from a cathode source. When these fields are suitably arranged, they have a focusing action on the electron beam; actually the field can act as a lens or a mirror. In the former case they produce a real 'cross-over' or minimum section of the beam within the lens field, and a cathode image further away. Moreover, they have a focal length which is necessarily small, and related to the distances of object and image by a slight modification of the usual lens formula.

The present study of these properties is mainly concerned with establishing the form of the field under various conditions of modulation. The resulting gradients at the cathode surface in a field which is space-charge free are worked out for the different cases. The effect of variations in geometry on the distribution are described since the information yielded has a direct bearing on the design of electron guns.

§ 2. METHOD OF EXPERIMENT

The field-distribution plots along and off the axis were obtained using an electrolytic tank in which distilled water was the conducting medium. Half-section ten-scale models of cathode, modulator and accelerator, which composed the objective, were silver-plated and suspended from two parallel ebonite rods attached to the removable tank-cover of brass. The probe was a silver wire, 0·2 mm. in diameter, fixed in a small Perspex chuck, which was rigidly connected to the movable screw-thread S. Figures 1(*a*) and 1(*b*) show the tank and electrode arrangement respectively. The set-up was similar to that described in an earlier paper (Jacob 1938) and the same precautions had to be strictly observed. These early tank plots were obtained by using a null method to trace out the equipotentials, which were then recorded by an electromagnetic pencil placed vertically above

---

* Now at University of Liverpool.

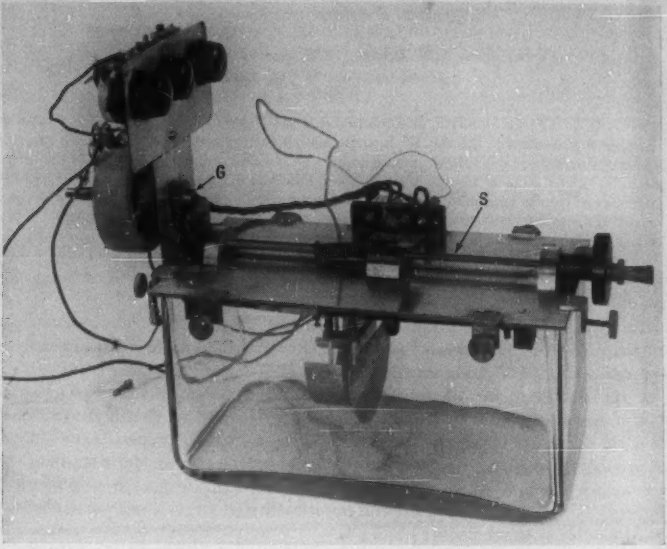


Figure 1 (*a*). Half-section ten-scale model in tank.

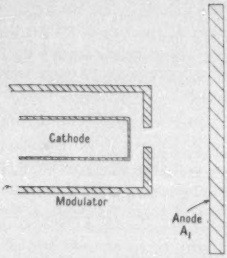

Cathode
Modulator
Anode
A₁

Figure 1 (*b*). Electron optical system.

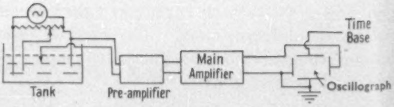

Time Base
Main Amplifier
Tank
Pre-amplifier
Oscillograph

Figure 2.

the probe. While the results obtained were accurate to 3%, the method was tedious in practice, requiring several hours for the complete plot. In the present case it was decided to work mainly on the axial distribution and to make the recording semi-automatic. This is achieved by picking up the potential difference between the probe and one electrode, applying it first to a pre-amplifier, and then to the main amplifier, whose output is fed to one pair of plates of an auxiliary oscillograph (Figure 2). The other pair of plates has a D.C. linear time-base potential, synchronized with the movement of the probe, through a system of gears G linked to the brass spindle S. In this way the potential distribution is seen as a fluorescent trace on the face of the cathode-ray tube.

Where a large number of traces are required—for example, when investigating the effect of variation of parameters such as cathode/modulator (C/M) spacing, anode/modulator(A/M) spacing, etc.—the method has proved expeditious, usually taking one minute per distribution. Adjustment to the electrodes was facilitated by removal of the brass cover top from which they were suspended.

One special feature of the present method was the ability to amplify any small portion of the path of the probe independently of the rest of the path. This was especially useful where it was desired to investigate the conditions obtaining close to the cathode when the slope of the field was small. The evaluation of the gradient then became practicable, since magnifications of ten or more could be used. Owing to the rectifying action of the pre-amplifier, the traces are, of course, recorded as D.C. on one side of the base-line only, and it is necessary in interpreting the plots to bear this in mind. The traces were photographed and enlarged for any subsequent calculations of gradients, emitting area, and so on.

§ 3. THE AXIAL POTENTIAL DISTRIBUTION

The chief region of interest in the axial plot lies between the cathode and some point up to or beyond the modulator for which the distribution of field is non-uniform (it was found in the preliminary work that between this point and the anode the potential rise was practically linear). It was hoped, by concentrating attention on this region, that a law might be found to enable the potential to be expressed as a function of distance from the cathode. Such a relation would simplify both calculations for the electron trajectory and those for establishing the cardinal points of the lens system. With this in view, families of curves were obtained for the axial distribution under the condition that the modulator was at cathode potential.

These are shown in Figure 3 for different C/M spacings, where the unit of distance is in terms of the modulator hole diameter, and the curve generally taken well beyond the position of the modulator. A plot of $\log V$ against distance yields a corresponding family of curves which are approximately straight lines, Figure 4, and consequently can be expressed in the form $V = Ae^{kz}$ over the limited region explored, where A and k are constants associated with the voltage and geometry respectively. The Table gives their values for the various C/M spacings for this type of system, where the modulator hole diameter is again taken as the unit of length.

C/M	0·05	0·10	0·20	0·30	0·50
A	0·056	0·066	0·069	0·058	0·018
k	1·61	1·30	0·83	0·74	0·78

In order to satisfy the boundary condition at the cathode ($V=0$), it is necessary to express the potential as a sinh function given by

$$V=A(e^{kz}-e^{-kz}) \qquad \ldots\ldots(1)$$

and assume that the value of k remains unchanged as the cathode is approached. This means that a uniform field must exist in the immediate vicinity of the cathode; this soon merges into the simple exponential field derived from the field plots, while further along the axis towards the anode the distribution becomes linear again.

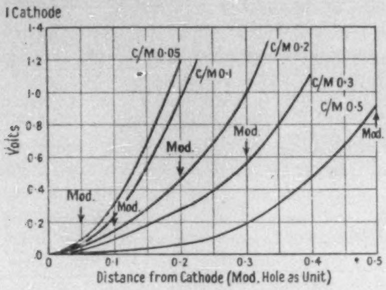


Figure 3.

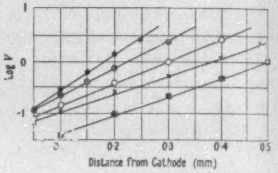


Figure 4.

● C/M 0·05 mm. : $V=0{\cdot}056\,e^{1{\cdot}61z}$
⊕ ” 0·10 mm. : $V=0{\cdot}066\,e^{1{\cdot}3z}$
○ ” 0·20 mm. : $V=0{\cdot}069\,e^{0{\cdot}83z}$
× ” 0·30 mm. : $V=0{\cdot}058\,e^{0{\cdot}74z}$
■ ” 0·50 mm. : $V=0{\cdot}0182\,e^{0{\cdot}78z}$

Such a form of distribution along the axis can be derived by considering the solution of the Laplace equation for rotationally symmetric systems:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right)+\frac{\partial^2 V}{\partial z^2}=0. \qquad \ldots\ldots(2)$$

When the potential V is put in the form $V=(\phi)\times(W)$, where ϕ is a function of r only and W is a function of z only, the solution for points on the axis takes the form

$$W=\Sigma\{A_n\exp(k_n z)+B_n\exp(-k_n z)\}. \qquad \ldots\ldots(3)$$

It thus might be possible, under certain conditions, to predict a form of exponential law as established empirically by the field plots.

In order to examine the early stages of the distribution in greater detail and to determine the effect of varying the modulator potential with respect to the cathode, use was made of the special amplifying facilities available. Three cases were investigated: (*a*) modulator at a positive potential; (*b*) modulator negative, up to that value which neutralizes the field from the anode—the cut-off or ‘black-out’ voltage V_B; (*c*) modulator negative, beyond the ‘black-out’ point. These are shown in Figures 5 (*a*), (*b*) and (*c*) respectively. At the ‘black-out’ point there is practically zero field at the centre of the cathode; the limiting cut-off current which is used in visual checking with sealed-off tubes must be made up of those electrons which, on account of their initial velocities, pass over the small negative barrier immediately in front of the cathode. Figure 5 (*c*) shows that beyond the ‘black-out’ point there is a reversal of curvature on the actual traces. This arises from the rectifying action of the amplifier, so that the potential distribution is recorded in one direction only.

In the diagram, the positive direction has been taken downwards. To avoid confusion, only one of the curves, e.g. $V_M = -4{\cdot}3$ v., is shown as it should appear. The negative minimum between the zero equipotential and the cathode indicates that the field has now taken the saddle-back form.

The exponential form of the region in the vicinity of the cathode can be deduced from Figure 6 for the different (negative) modulation voltages applied.

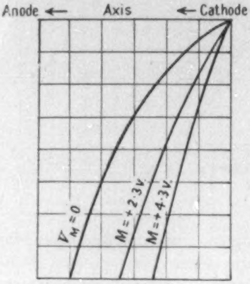


Figure 5 (*a*).

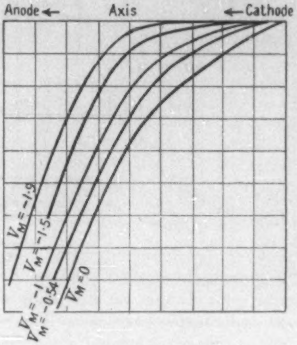


Figure 5 (*b*).

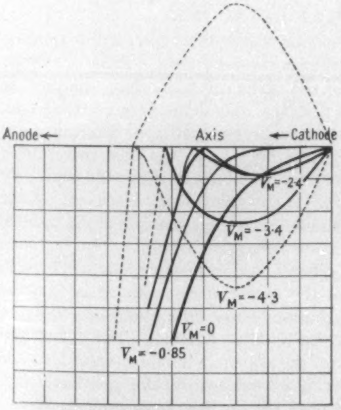


Figure 5 (*c*).

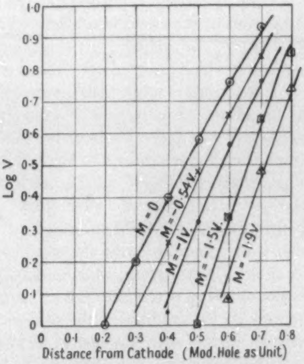


Figure 6.

Here again the curves are approximately straight lines with a slope increasing gradually as the modulator becomes more and more negative towards and past the 'black-out' point. The gradient at any point on the axis is thus proportional to the potential at that point, at least for static conditions. At the cathode surface the gradient is, of course, not zero, except in very special cases. Its value is $2kA$, obtained by differentiating equation (1) with respect to z. An analysis of the curves of Figures 5 (*a*) and 5 (*b*) shows that the linear law holds only in cases of positive gradient all the way to the cathode. Where, as in Figure 5 (*c*), the field reverses in the neighbourhood of the modulator, the relationship becomes somewhat complicated.

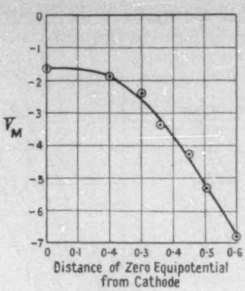


Figure 7.

The potential at any fixed point on the axis decreases with decrease in modulator potential; on the other hand, a point at a fixed potential moves away from the cathode under these conditions. This is illustrated in Figure 7, which shows the shift of the zero equipotential along the axis as the potential of the modulator is decreased; the effect is most marked for the early stages of modulation.

§ 4. GRADIENT AT CATHODE

Two important experimental relationships which affect the current-emitting properties of the cathode can be deduced from the field plots. These are (i) the gradient at the cathode, (ii) the point in the radial direction along the cathode surface at which the gradient normal to it vanishes; this defines the emitting area.

It was usual to work with the modulator at cathode potential for determinations of the slope. Let this be S volts/cm. The modulator was then biased back negatively to a value V_B in order to neutralize the effects of the anode field. By combining these values, the slope *in vacuo* for any value of the black-out voltage V_M is then $10S(V_M/V_B)$. Figure 8 (*a*) shows the average gradient at the cathode at $V_M=0$, calculated for a constant black-out voltage of 30 volts. This has been plotted against the C/M spacing for two different hole diameters. It will be seen that over a certain range the gradient at the centre of the cathode is constant at about 200 volts/cm. independent of the modulator hole diameter. For the larger spacings it tends to fall off rapidly for both modulators. The dotted line represents the gradient at the mid-point of the system. This increases approximately linearly with increase in C/M spacing for constant black-out voltage.

It is clear that the gradient will depend on the black-out voltage alone, provided that the anode voltage and geometry remain fixed in any one case.

The fact that kA is constant at the centre of the cathode for constant black-out voltage indicates the existence of a uniform field in this region. The 'cross-over' or minimum cross section of the beam is formed further away—in the exponential part of the field which extends to the other side of the modulating electrode, and which appears to govern the focusing properties of the lens.

The decrease in gradient at constant anode voltage is shown in Figure 8 (*b*). The ratio of gradients is in proportion to the square of the modulator hole diameter. As expected, the thinner modulator gives the larger gradient, but there does not appear to be any simple relation involving the thickness of the modulator alone. An interesting curve is that shown in Figure 9, where the ratio of anode to modulator voltage, V_A/V_M, for zero gradient at the centre of the cathode (i.e. for the black-out condition) is plotted against C/M spacing. The result is a linear relationship, as might have been predicted from orthodox theory. This ratio varies approximately inversely as the square of the hole diameter. From this we may infer that at constant anode voltage the gradient at the cathode surface is proportional to the modulator potential, and hence varies as the black-out voltage.

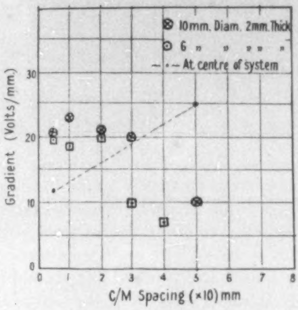


Figure 8 (*a*). Average gradient at cathode for V_{BO}=30 volts.

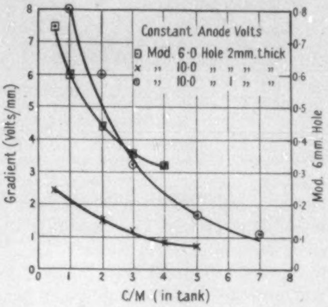


Figure 8 (*b*). Average gradient at cathode at M=0.

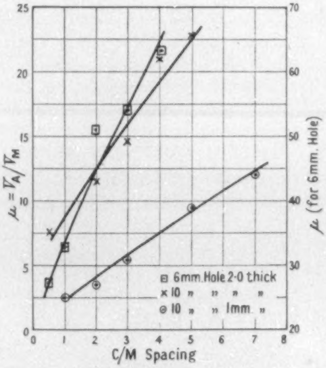


Figure 9.

A useful curve for design purposes is that connecting the black-out voltage V_B with C/M spacing, since the current yield can be calculated from the value of V_B. This is shown in Figure 10(*a*) and (*b*) for the 10 mm. and 6 mm. modulator holes respectively. It will be noticed that the rate at which the black-out voltage falls is greatly increased at the short spacings; this necessitates special care in assembly to avoid a large spread in V_B. In order to check the accuracy of the static determination in the tank, a number of sealed-off gun units were made up to one-tenth the tank scale. These were pumped and processed in the normal way, and the black-out point was determined in each case by the disappearance of the fluorescent spot on the screen. The values were then reduced so as to make only one point correspond to a point on the curve; the remaining points are seen to lie fairly close to the tank curve.

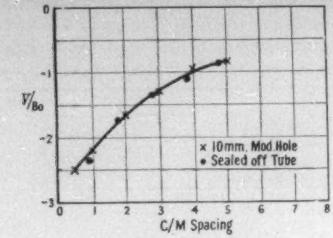


Figure 10 (*a*).

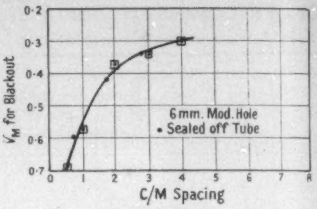


Figure 10 (*b*).

In general, the conclusions reached from a study of the results on the 10 mm. aperture modulator could be applied to the 6 mm. aperture; for example, the exponential character of the law connecting potential and distance from the cathode was upheld in all cases of C/M variations, e.g. the shift of the zero equipotential along the axis with modulation.

§ 5. THE EMITTING AREA

It was possible to explore the gradient along the cathode surface at points off its centre by moving the probe along it. The point at which the gradient normal to the surface becomes zero defines the maximum emitting area of the cathode. This is shown in Figure 11 for the case when the modulator is at cathode potential for each of two separate C/M variations.

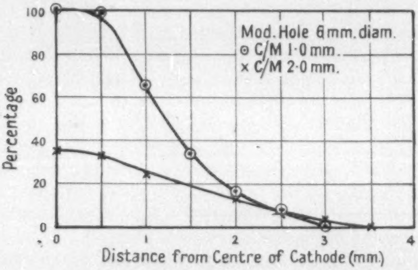


Figure 11. Average gradient as percentage of gradient at centre.

The ordinates when plotted as a percentage of the gradient at the centre of the cathode show a markedly increased rate of fall for the smaller spacing. The emitting area will be seen to correspond approximately with the projection of the modulator hole on to the cathode surface when, as in this case, the modulator electrode is tied to the cathode potential. For negative bias values the emitting area is proportional to the ratio of bias voltage to black-out voltage.

REFERENCES

Jacob, L., 1938, *Phil. Mag.*, **26,** 570.
Johannson, H., 1933, *Ann. Phys., Lpz.*, **18,** 385.
Maloff, I. G., and Epstein, D. W., 1938, *Electron Optics in Television* (New York and London : McGraw-Hill).
Moss, H., 1945, *J. Brit. Instn. Radio Engrs.*, **58,** 10,

Colour Perception in Parafoveal Vision

By M. GILBERT *
Imperial College of Science and Technology, London S.W.7

MS. received 1*st July* 1948, *and in amended form* 16*th August* 1949

ABSTRACT. Quantitative studies of extra-foveal colour vision are inherently difficult owing to the poor discrimination and rapid adaptation found in the parafoveal and peripheral areas of the retina. Measurements of the luminosity, colour mixture and hue discrimination curves have been made by a single observer using a field subtending 1° 20′ and located at 2° and 4° from the fovea. The results suggest that for high retinal illumination parafoveal colour vision does not differ in essential characteristics from foveal colour vision, but at low illuminations colour perception becomes reduced to a type of dichromatic vision similar to the tritanopic form of defective colour vision and to that reported for small-field foveal vision.

§ 1. INTRODUCTION

The most exhaustive studies of colour perception have in the past been made on foveal vision using fields subtending one or two degrees. Recently interest has been directed to the study of variations in colour vision across the retina, and experiments with quite small fields, of the order of 10–15 minutes of arc, have been carried out in an attempt to analyse the colour mixture and colour discrimination functions under conditions in which a relatively small number of receptors are stimulated. Interesting results (Willmer and Wright 1945, Thomson and Wright 1946, Stiles 1949) have been reported on variations of colour sensitivity even within that area of the fovea which it is customary to regard as being free from rods, and an extension of these investigations beyond the foveal boundary into the mixed rod–cone regions of the retina is obviously desirable. From such experiments we might hope to obtain a clearer idea of the manner in which the rod and cone processes interact and also of the differences, if any, between foveal and parafoveal cone sensitivity. Moreover, under everyday conditions of seeing, quite large areas of colour are normally in view, so that a study of parafoveal colour vision should have a direct bearing on practical colour problems.

---

* Now at Building Research Station, Garston, Watford, Herts.

No very great reliance can be placed on experiments in which the appearance of a colour seen by parafoveal or peripheral vision is described verbally and compared with its foveal appearance. For certain purposes such comparisons may be necessary and valid, but more quantitative data are to be preferred. The aim here has been to measure the luminosity, colour mixture and hue discrimination curves for parafoveal vision by procedures essentially similar to those which have been used in the past for foveal vision.

It has proved very difficult to obtain results having a precision and accuracy sufficiently high to yield a significant measure of some of the variations of colour perception that occur across the retina. The main reasons for this appear to be the drop in visual acuity, the reduced ability to discriminate luminosity and colour differences, and the greater and more rapid adaptation in parafoveal vision compared with foveal vision. The poorer visual acuity implies that if the two halves of a colorimeter field are made too small, they will no longer be resolved in the parafovea, in which case colour matching and colour discrimination become impossible. On the other hand, if the field were made too large, the possibility of exploring individual areas of the retina would be lost, since a field subtending 6°, for example, and placed with its centre located at a point 3° to one side of the foveal fixation target, would evidently stimulate a large fraction of the fovea in addition to the parafovea. The retinal structures viewing the field would then vary continuously and markedly along the horizontal diameter of the field.

In the present experiments, a square field subtending 1° 20′ at the eye was used, this size enabling the two halves of the field to be resolved while at the same time retaining fairly local stimulation of the retina. With this size of field, however, discrimination of both luminosity and colour differences was not good, which in turn meant that precision of matching was poor. This was undoubtedly due in part to the smallness of the field and also to the inherent difficulty of making observations on a part of the field at which the observer is not looking directly.

The greater adaptability of the parafoveal retina also tends to lower the accuracy of observation, since different stimuli may give rise to rather similar sensations owing to compensatory adjustments in the sensitivity of the retinal areas used to view them. It can even happen that when prolonged and steady fixation is maintained on the fixation target, so that the parafoveal field under investigation is illuminating the same part of the retina for some time, the field may almost, if not quite, disappear. Fixation for more than a few seconds at a time has, therefore, to be avoided.

§ 2. APPARATUS AND OBSERVING CONDITIONS

The apparatus used in the experiments was the Wright colorimeter (Wright 1946). The observer's head was fixed by biting on a dental impression and was positioned so as to bring the exit-pupil of the colorimeter in line with the axis of the observer's eye. The field subtended 1° 20′ at the eye and was divided horizontally into two equal parts. The upper half, the matching field, was illuminated by any or all of three monochromatic radiations, of wavelengths $0{\cdot}65\ \mu$, $0{\cdot}53\ \mu$ and $0{\cdot}46\ \mu$, or, if required, by monochromatic light of any wavelength through the spectrum, while the lower half, the test field, was illuminated by either one or two monochromatic radiations of any desired wavelengths. A small red fixation spot, displaced horizontally from the test field, enabled the latter to be imaged on the required part of the parafovea when the fixation spot was viewed foveally. Provision was also made for the test field to be separated from the matching field by a few degrees to permit direct comparison of the sensitivity of one part of the retina with another. The adaptation system on the colorimeter was available to condition the eye to various degrees of light and colour adaptation.

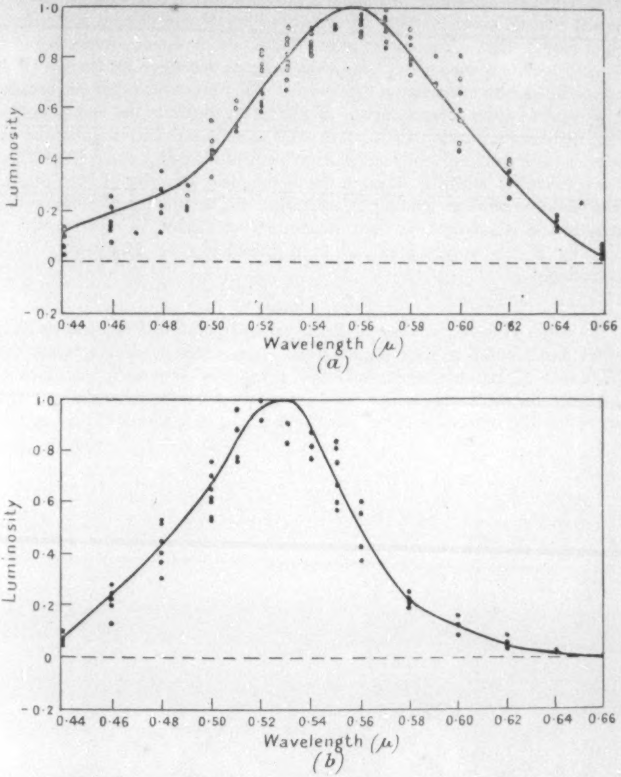


Figure 1. Luminosity curves for the parafoveal retina at a position 4° from the foveal centre, using linear luminosity scale on the ordinates. The magnitude of the errors in making the observations is indicated by the spread of the individual observed points shown in the diagram. The curves have been drawn through the arithmetic means of the observed points and the relative luminosity values have been adjusted so that the maximum of each curve is always unity.

(*a*) Curve measured with the eye adapted to white light (brightness of the order of 1,000 photons, colour temperature 2,848° K. approximately) and with a test field of a similar brightness.

(*b*) Curve measured with the eye dark adapted, and with a test field brightness of approximately 2·2 photons.

Unless otherwise indicated, the measurement and calculation of the luminosity curves, colour mixture curves and hue discrimination curves were carried out essentially as described by Wright (1946) for foveal vision.

§ 3. RESULTS

Luminosity

Typical luminosity curves for the parafoveal retina at a position 4° from the foveal centre are illustrated in Figure 1, curve (*a*) being recorded for the light-adapted eye and curve (*b*) with the eye dark-adapted and using a relatively low retinal illumination. The curves were determined by direct comparison of the test field against a monochromatic wavelength, as described by Wright (1946). Several comparison wavelengths, e.g. 0·46 μ, 0·65 μ, were used on different occasions for repeat observations of each curve. These curves illustrate the commencement of the well-known Purkinje shift, but in some cases humps and irregularities have been recorded, more especially in the neighbourhood of 0·51 μ and 0·56 μ. These have presumably had their origin in the simultaneous activity of both rods and cones under conditions tending to accentuate the sensitivity of either rods or cones at the wavelength of their maximum sensitivity. Unfortunately, the reliability of the observations has been insufficient to analyse the humps convincingly.

Colour Matching

No major differences between the foveal and parafoveal spectral coefficient curves were found at high illumination. Measurement of the white point (illuminant S_B, colour temperature 4,800° K.) showed systematic variations with distance from the fovea, as indicated by curve *a* in the chromaticity chart, Figure 2. The movement of the point towards the blue corner of the chart with

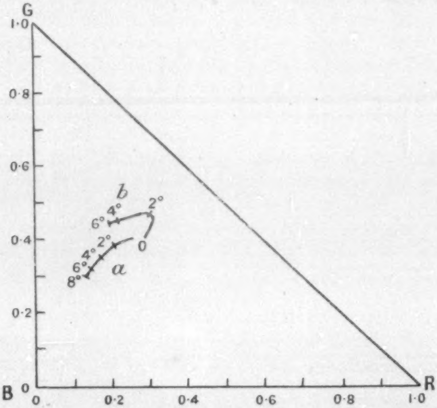


Figure 2. Chromaticity chart showing variations of the white point (standard illuminant S_B, colour temperature 4,800° K.) with distance from the fovea. Eye adapted to light of approximate colour temperature 2,848° K., and brightness 1,000 photons. The wavelengths at the points R, G and B are 0·65 μ, 0·53 μ and 0·46 μ respectively, and the units marked along the side of the triangle refer to the unit trichromatic coordinates. Observations were made with:

(*a*) Test and comparison fields both observed in the parafovea by directing the gaze at a small red fixation point. The figures 0° to 8° on the curve *a* indicate the coordinates measured with both fields at those distances from the fovea.

(*b*) Test field viewed in the parafovea whilst the comparison field is viewed foveally. The figures 0° to 6° on the curve indicate the coordinates measured with the test field at those distances from the comparison field at the fovea.

increasing distance from the fovea may correspond to the diminishing density of the yellow macular pigment, but this interpretation is complicated by the simultaneous variation of both retinal structure and macular pigmentation.

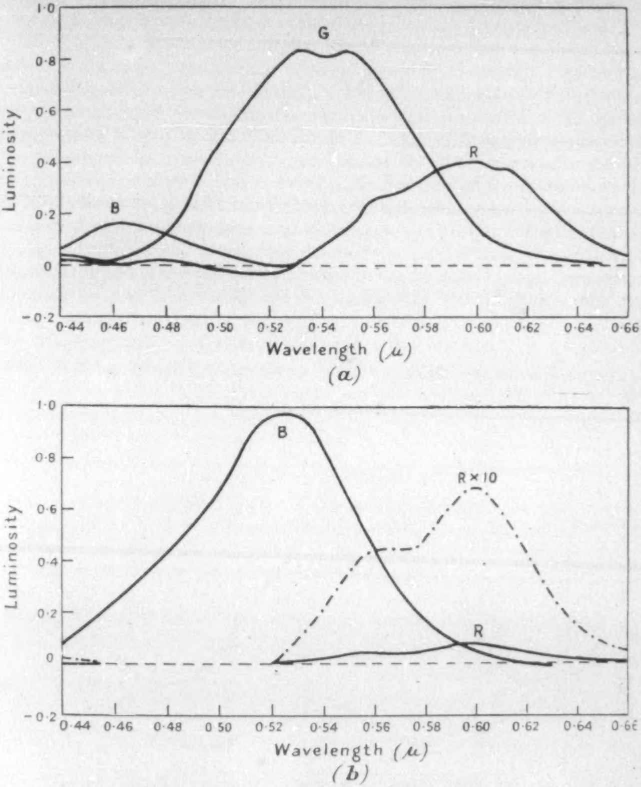


Figure 3. Colour mixture curves for the parafoveal retina at a position 4° from the foveal centre, using linear luminosity scale on the ordinates. These curves have been derived from the corresponding luminosity curves of Figure 1 and measurements of trichromatic or dichromatic coefficients, the calculations being made as described by Wright (1946).

(*a*) Measurements made with the eye adapted to white light (brightness of the order of 1,000 photons, colour temperature 2,848° K. approximately) and with a test field of the same order of brightness.

(*b*) Measurements made with the eye dark-adapted, and with a test field brightness of approximately 2·2 photons.

When colour matches were made between the test field viewed at various positions in the parafovea and the red-green-blue mixture viewed foveally, it was found that relatively less red and more green was required to match a monochromatic yellow test colour, as the test field was displaced parafoveally.

This difference can only be due to differences in the visual processes and not to pigmentary absorption, since the latter cannot cause any change in the quality of monochromatic light. The corresponding effect on the white point is shown by curve *b* in Figure 2.

At low illuminations and with both fields viewed parafoveally, colour matching tended to become dichromatic. All the spectral colours, for example, could be matched by a mixture of only two stimuli, one of these being a red radiation (for example 0·65 μ) and the other either a green, blue-green or blue (for example 0·46 μ). The dichromatic mixture curves obtained at 4° displacement into the parafovea are illustrated in Figure 3 (*b*). Two points of special interest are the small contribution made by the red mixture curve to the total luminosity curve, and the hump in the red curve itself. There is such a striking similarity to the red mixture curve found by Willmer and Wright (1945) for small field foveal vision, that the hump must almost certainly have a common origin in the two cases. The trichromatic mixture curves obtained at 4° displacement into the parafovea when the eye is light adapted are illustrated in Figure 3 (*a*). Both the red and green curves show humps in this case.

Hue Discrimination

The most noticeable difference between foveal and parafoveal hue discrimination at high illuminations was the larger wavelength difference required in parafoveal vision to produce a just noticeable difference in hue. A typical curve for 4° parafoveal viewing is shown in Figure 4, curve *a*, in which maximum sensitivity (minimum wavelength difference) occurs in the orange and blue-green parts of the spectrum, as with foveal vision.

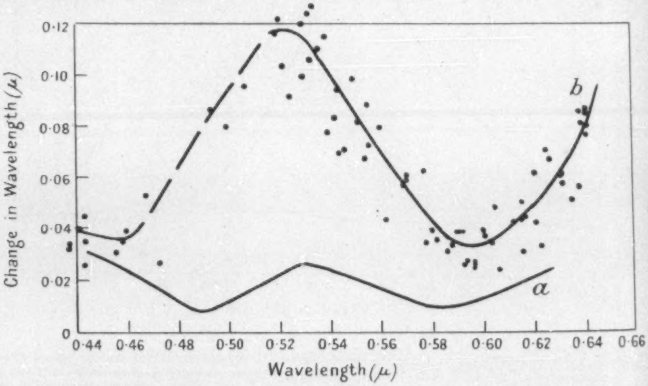


Figure 4. Hue discrimination curves for the parafoveal retina at 4° from the foveal centre. Individual observed points give an indication of the errors of observation involved.

(*a*) Eye adapted to light of brightness approximately 1,000 photons, colour temperature 2,848° K., and test field of the same order of brightness.

(*b*) Eye dark-adapted and test field brightness approximately 1·1 photon.

At low illuminations, however, there is a definite alteration in the type of discrimination curve, since discrimination in the blue-green wavelengths becomes very poor indeed. Curve *b* of Figure 4 consists of a single minimum in the orange, with some isolated points in the blue and violet. There is, in effect, a discontinuity in the blue-green, similar to that reported by Willmer and Wright (1945) for small field foveal vision, since the blue and green wavelengths are virtually indistinguishable.

§ 4. CONCLUSIONS

Many of the experimental data, more especially those relating to the effect of adaptation of parafoveal colour vision, have been omitted from this report because of lack of precision of the measurements. The first conclusion which must be drawn therefore is the inherent difficulty of securing results on parafoveal colour vision which can have any claim to useful accuracy. From the more conclusive of the results which have been recorded it appears (i) that parafoveal colour vision, at high illuminations does not differ in essential characteristics from foveal vision, although secondary differences do exist, and (ii) that at low illuminations, colour discrimination becomes less acute and is reduced to a form of dichromatic vision which is rather similar to the tritanopic class of defective colour vision and also to that reported for small field foveal vision.

One explanation given for this type of reduced discrimination is inactivity of the blue receptors. Alternatively, it may be that the presence of rods in the parafovea adds a colourless quality to the blue and blue-green response at low illuminations, which swamps the blueness and greenness, whereas at the orange and red end of the spectrum, the rods are relatively much less sensitive than the cones and little or no dilution of hue can occur.

ACKNOWLEDGMENTS

This work was carried out as part of a research programme under the direction of Professor W. D. Wright; his constant help and advice during the observations and in preparing this paper are greatly appreciated.

My thanks also go to Misses M. E. Radley and M. J. Thornley, Dr. L. C. Thomson and Mr. J. H. Dodwell for help in recording the observations, and to the Medical Research Council for their financial support of the investigation.

REFERENCES

STILES, W. S., 1949, *Rev. d'Optique*, **28,** 215.
THOMSON, L. C., and WRIGHT, W. D., 1946, *J. Physiol.*, **105,** 316.
WILLMER, E. N., and WRIGHT, W. D., 1945, *Nature, Lond.*, **156,** 119.
WRIGHT, W. D., 1946, *Researches in Normal and Defective Colour Vision* (London: Kimpton).

Quantum Efficiency in Photographic X-ray Exposures*

D. BROMLEY† AND R. H. HERZ
Research Laboratories, Kodak Ltd., Wealdstone, Harrow, Middlesex

MS. received 30*th May* 1949, *and in amended form* 4*th August* 1949

ABSTRACT. The grain yield per absorbed quantum has been evaluated experimentally for four different photographic emulsions over a wavelength range of 1·2–0·01 A., i.e. from very soft x rays to gamma rays emitted from radium. The tests were carried out with both heterogeneous and monochromatic x rays. A grain yield of 1·0 per absorbed quantum appears to hold within a range of wavelengths of 1·2–0·37 A. (10 kv.–33 kv.). From 0·37 A. to short wavelengths (gamma rays) an increase of grain yield linear with quantum energy is obtained. At a wavelength of 0·0123 A. (1,000 kv.) about 80 grains are rendered developable per quantum absorbed. The rate of increase of grain yield with quantum energy is different for different emulsions. For a slow and fine-grain chloro-bromide emulsion the grain yield was found to be less than unity for wavelengths below 0·12 A. (i.e. 100 kv.). The results are consistent with the hypothesis that grains are made developable not only by the direct action of x rays but also by secondary electrons produced by the x rays. These are effective only when they are energetic enough to reach grains in the neighbourhood of that absorbing the original x-ray quantum.

§ 1. INTRODUCTION

In order to understand the mechanism of direct photographic x-ray exposures a knowledge of the efficiency of the x-ray energy absorbed in photographic emulsions as a function of the wavelength is required. This efficiency may be expressed as the number of silver bromide grains rendered developable per absorbed quantum, often called the quantum efficiency or the grain yield per absorbed quantum.

Eggert and Noddack (1927) determined the grain yield per absorbed quantum of x rays in photographic emulsions and found that each quantum absorbed renders approximately one silver bromide grain developable. They used a heterogeneous radiation equivalent to a monochromatic radiation of wavelength 0·45 A. The determination of this equivalent wavelength was based on calorimetric energy measurements. Silberstein and Trivelli (1930) gave somewhat less direct evidence that the one quantum hypothesis holds in the range of x-ray wavelengths. Similar investigations to those by Eggert and Noddack were carried out by Günther and Tittel (1933) within the range of wavelengths from 0·245 A. to 1·54 A., which agreed well with the results of the former workers. In addition they found that the number of silver atoms set free is proportional to the energy of the x-ray quantum. Pelc (1945) has recently given a theoretical and experimental account of the increase of grain yield with increasing quantum energy.

The purpose of the work described in this paper is to examine the quantum efficiency of various types of photographic emulsions over a wide range of x-ray quality, by providing special experimental techniques in evaluating the grain yield. These techniques consisted in measuring the actual absorption of x rays by emulsions stripped from film support, and in using in some cases monochromatic x rays instead of heterogeneous radiation. By this latter procedure a greater absolute accuracy in the determination of the grain yield per absorbed quantum was expected.

---

* Communication No. 1249H from the Kodak Research Laboratories.
† Now with the Royal Naval Scientific Service.

§2. EXPERIMENTAL METHODS AND CALCULATIONS

Experiments were carried out in the first place by exposing films to measured doses of filtered heterogeneous X rays whose quality was determined in terms of half-value layers in copper by means of an ionometric half-value layer meter. The half-value layer values were converted into absorption coefficients and the equivalent wavelength of the radiation was then determined as the wavelength of that monochromatic radiation having the same absorption coefficient.

In four cases monochromatic X rays were used from secondary radiators, such as platinum, copper and barium carbonate, which acted as characteristic X-ray sources when irradiated by heterogeneous X rays from a tungsten target tube. The K_β radiations were suitably filtered by selective absorption filters. This method of obtaining monochromatic X rays was regarded as preferable to the use of a monochromator because it provides a fairly wide and uniform beam, suitable for density and ionization measurements.

The dose was measured by means of a Victoreen condenser r-meter, or, in the case of very soft X rays, by means of another ionization dosemeter using an ionization chamber with 'Cellophane' window. Both meters were calibrated against a free air standard ionization chamber at the National Physical Laboratory. The number of X-ray quanta, Q_i, incident upon the emulsion for various wavelengths was computed from the definition of the roentgen dose. The number of quanta absorbed (Q_a) per unit area in the photographic emulsion was then calculated as $Q_i(\mu/\rho_{\mathrm{AgBr}})w$, where μ/ρ_{AgBr} is the mass absorption coefficient of silver bromide, and w the weight of the silver bromide in gm/cm^2 of emulsion. Details of this computation are described below.

The number of developed grains in the emulsion was then counted under the microscope, using a squared graticule and utilizing only that emulsion layer (in the case of double coated X-ray films) facing the X-ray tube during exposure. The number of fog grains, counted on unexposed and developed film strips, was subtracted from the former number. Since the counting of grains under the microscope is tedious and the statistical error involved rather great, the number of grains rendered developable was also evaluated from density measurements, assuming that density is proportional to the number of grains affected. These tests were limited to the range where the relationship between density and exposure is linear.

The desired number of developed grains per quantum absorbed in the emulsion was found from the ratio of the number of grains counted per unit area to the number of quanta absorbed per unit area.

The energy of X rays absorbed E_{abs} within a small mass of air dm may be calculated from the product of the energy flux E_{fl} in ergs/cm^2, the mass in grammes of the air, and the fraction of the energy flux, incident upon the mass, that is converted into kinetic energy of electrons. This fraction is given by $(\tau+\sigma_a)_{\mathrm{air}}$, where τ is the photoelectric absorption coefficient and σ_a the true absorption coefficient associated with Compton recoil electrons. Thus

$$E_{\mathrm{abs}} = E_{\mathrm{fl}}\, dm(\tau+\sigma_a)_{\mathrm{air}}. \qquad \ldots\ldots(1)$$

Hence

$$E_{\mathrm{fl}} = \frac{E_{\mathrm{abs}}}{dm(\tau+\sigma_a)_{\mathrm{air}}}. \qquad \ldots\ldots(1\,a)$$

By definition, one roentgen unit (r-unit) produces 1 E.S.U. of ions of each sign in 1 cm^3 of air. The number of electrons collected in 1 cm^3 (=0·001293 gm.) of air per roentgen is $1/\boldsymbol{e}$, where $\boldsymbol{e}$ is the electronic charge in E.S.U. The energy required to produce one pair of ions in air is approximately 33 electron volts. Thus the energy absorbed per cm³ of air for 1 roentgen dose is $33/\mathbf{e}$ electron volts. Since 1 ev. $=\mathbf{e}/300$ ergs, the dose of 1 roentgen corresponds to the absorption of $33/300=0{\cdot}11$ ergs. Since 1 roentgen corresponds to the absorption of 0·11 erg in 1 cm³ of air, equation (1 *a*) may be written

$$E'_{\mathrm{fl}}=\frac{0{\cdot}11}{dm(\tau+\sigma_{\mathrm{a}})_{\mathrm{air}}}\,\mathrm{erg/cm^2}, \qquad \ldots\ldots(1\,b)$$

where E'_{fl} is now the energy flux per roentgen. This expression is a function of wavelength through the term $(\tau+\sigma_{\mathrm{a}})_{\mathrm{air}}$. The variation with wavelength of the energy flux to give 1 roentgen in air has been computed and is shown graphically in Figure 1. In accordance with similar computations by Mayneord (1940) the curve possesses a maximum at about 0·1 A. where the energy flux is more than six times as great as at 0·4 A.

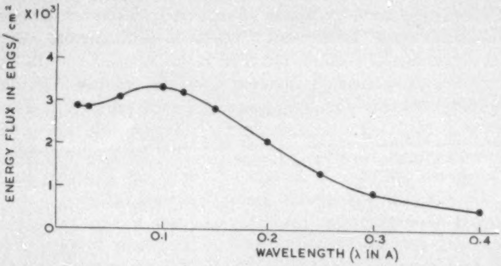


Figure 1. Relationship between energy flux in ergs/cm² and wavelength for 1 r-unit.

The energy absorbed, at a given wavelength, per cm² of a thin layer of silver bromide emulsion can be computed for any dose by multiplying expression (1 *b*) by the mass absorption coefficient of silver bromide, by *w*, the coating weight per cm² of the emulsion, and by the number of roentgens, *r*. Thus

$$E_{\mathrm{abs}}=\frac{0{\cdot}11\mu wr}{(\tau+\sigma_{\mathrm{a}})_{\mathrm{air}}\,\rho}\,\mathrm{erg/cm^2}. \qquad \ldots\ldots(2)$$

The gelatin absorption is neglected in these considerations as calculations have shown that the error involved is less than 1 %. The energy flux E_{fl} for one roentgen unit may be converted into the number of quanta Q by dividing it by the energy of the quantum concerned, thus

$$Q=\frac{E_{\mathrm{fl}}}{\boldsymbol{h}\nu}=\frac{\lambda E_{\mathrm{fl}}}{\boldsymbol{hc}}=\frac{0{\cdot}11\lambda}{\boldsymbol{hc}(\tau+\sigma_{\mathrm{a}})_{\mathrm{air}}}=\frac{\lambda}{(\tau+\sigma_{\mathrm{a}})_{\mathrm{air}}}5{\cdot}53\times10^{14}.$$

If λ is expressed in Ångström units

$$Q=\frac{\lambda}{(\tau+\sigma_{\mathrm{a}})_{\mathrm{air}}}5{\cdot}53+10^{6}. \qquad \ldots\ldots(3)$$

The number of quanta absorbed per roentgen unit in 1 square centimetre of a silver bromide emulsion is then obtained by multiplying the expression in equation (3) by the combined mass absorption coefficient of silver bromide and the weight of the silver bromide per cm² of emulsion (neglecting the gelatin). If the density of AgBr is taken as 6·47, then

$$Q_{\mathrm{abs}} = \frac{\mu w \lambda}{(\tau + \sigma_a)_{\mathrm{air}}} 0{\cdot}855 \times 10^6. \qquad \ldots\ldots(4)$$

The combined mass absorption coefficient of silver bromide for various wavelengths was found according to

$$\mu_{\mathrm{AgBr}}/\rho = \frac{w_1}{w_1 + w_2} \cdot \frac{\mu}{\rho_{\mathrm{Ag}}} + \frac{w_2}{w_1 + w_2} \cdot \frac{\mu}{\rho_{\mathrm{Br}}},$$

where w_1 and w_2 are the weights of silver and bromine respectively.

§ 3. SELECTION AND DETERMINATION OF WAVELENGTHS

(i) *Heterogeneous Radiation*

In order to cover a wide range of wavelengths in the first instance heterogeneous radiations were used. The generating kilovoltage, filtration and the equivalent wavelengths of these radiations are given in Table 1. The equivalent wavelengths were derived from half-value layer measurements partly on copper and partly on silver.

Table 1

kv.	Filtration	Equivalent λ (A.)
50	0·4 mm. Cu	0·30
80	0·8 mm. Cu	0·21
120	1·2 mm. Cu	0·16
195	4·0 mm. Cu	0·10
475	4·53 mm. Sn	0·047
1000	2 mm. Al + 4·2 mm. Fe + 6·5 mm. Pb	0·024
Gamma rays	1·0 cm. Pb	0·0105

The results on quantum efficiency which are given in § 5, are based on the wavelengths tabulated above.

As there was some doubt in regard to the accuracy of the wavelength determination of heterogeneous radiations which might seriously affect the results on quantum efficiency, the wavelengths were checked indirectly by absorption measurements with stripped silver bromide emulsions. The results of these were then compared with those obtained from computation using the mass absorption coefficients for the wavelengths from half-value layer measurements.

The absorption measurements were made by means of an X-ray ionization dosemeter comprising a Lindemann electrometer and two thimble ionization chambers of identical design and equal volume. The central electrodes of the two thimble chambers were connected to the needle of the electrometer. The walls of the chamber were connected to equal and opposite potentials (54 v.). If the volumes of the chambers are equal and if the chambers are exposed to a common uniform X-ray beam, so that both chambers receive equal doses of X rays, the electrometer needle will be in electrical balance. If a stripped photographic emulsion having no support is placed above one of the chambers the electrometer needle will be out of balance by an amount proportional to the absorption in the film which may thus be measured. The percentage absorption is given by $100(I_0 - I_1)/I_0$, where I_0 is the intensity of incident radiation, and I_1 the intensity transmitted by the emulsion. The method described measures $I_0 - I_1$ directly, leaving the separate values of I_0 and I_1 unknown. The value of

I_0 was determined simultaneously by means of the Victoreen condenser r-meter. The method is thus substantially independent of fluctuations in the beam, whereas in measuring such small absorptions by successive readings with and without absorber, errors in the meter readings and intensity differences due to fluctuations of the intensity of the beam are generally of the same order as those due to absorption. In actual use without the absorber in position, an exact balance is not easily obtained, but a small residual drift remains. This drift, with appropriate sign, was therefore subtracted from the deflection due to absorption as actually measured.

Absorption measurements were made with 1, 2, 3 .... superimposed stripped emulsions placed about 1 inch above one of the chambers. The results obtained, which were plotted as percentage absorption against the number of stripped emulsions, showed more than a proportional increase in absorption with the number of sheets. The effect was ascribed to scatter and fluorescent x rays from the absorbing sheets. The influence of scatter was however eliminated by placing the absorbing sheets 8 in. above the measuring chamber.

A typical result of such an absorption measurement using a stripped emulsion without base is shown in Figure 2. The absorption of one sheet may be read from the graph and is 0·37% for radiation the wavelength of which was computed from half-value layer measurements to be 0·21 A. The percentage absorption to which this wavelength leads using the absorption coefficients from tables is in this particular case exactly the same. It may be mentioned that the film used was a chloro-bromide emulsion and that the constituents of this emulsion were considered in calculating the absorption for various wavelengths. Absorption measurements were also carried out on another type of stripped emulsion using radiations generated at 50 kv. filtered by 0·4 mm. Cu. at 80 kv., (0·8 mm. Cu), and at 120 kv. (1·2 mm. Cu).

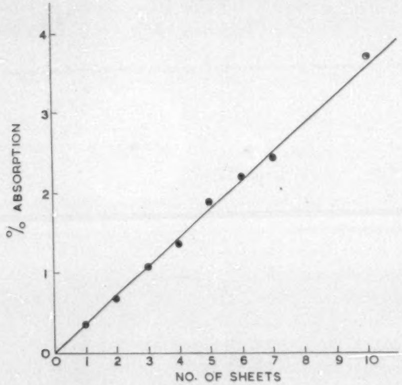


Figure 2. Relationship between percentage absorption and number of stripped emulsion layers for x-rays generated at 80 kv., 8 ma., 0·8 mm. Cu-filter (equivalent wavelength $\lambda = 0{\cdot}21$ A.)

In addition half-value layer measurements of these radiations in silver and copper were made and, using the equivalent wavelengths, the percentage absorption was evaluated. The calculated and the experimentally obtained percentage absorption values agree fairly satisfactorily. The greatest deviation occurring between two differently evaluated wavelengths was not more than 7%.

It was mentioned above that half-value layer figures were also determined in terms of millimetres of silver. This was done as it was found that the effective wavelength of the same heterogeneous radiation differed in different material. Since the absorption in the silver bromide emulsion is predominantly that of silver, it was thought that the half-value layer in silver would give a better estimate of the actual quality of x rays, so far as the emulsion is concerned. The results show however only small differences between the half-value layers obtained in copper and in silver.

(ii) *Monochromatic Radiation*

In view of the uncertainty of the wavelengths used as computed from heterogeneous radiations, experiments were carried out with the object of producing monochromatic radiations The methods available are (1) reflection from a crystal, (2) use of secondary characteristic radiation excited by the absorption of primary x rays. Method (1) gives a sufficiently intense but narrow beam of x rays although any particular line of known wavelength may be selected or any wavelength may be calculated using the Bragg relation. Ionization measurements in terms of r-units become difficult, however, because of the narrow width of the beam. Method (2) gives a broad beam of low intensity. Careful screening from the primary beam is necessary. The spectrum normally comprises several lines together with scattered and continuous radiation. The latter method appeared to be preferable for this investigation chiefly because preliminary tests have shown that a wide and relatively uniform x-ray beam could be obtained, which allowed film exposure and ionization measurements to be made simultaneously*. The experimental set up is illustrated in Figure 3. The apparatus consists mainly of a large lead lined box. One side of this is removed, and to the side

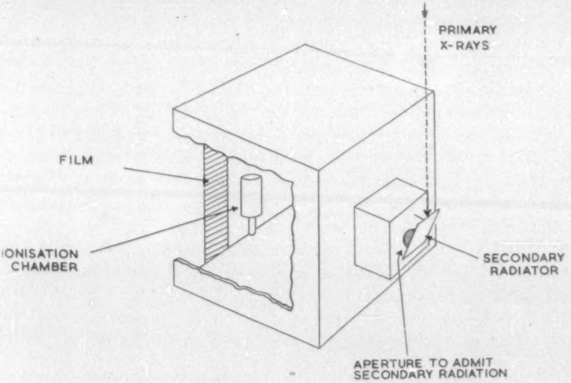


Figure 3. Arrangement for generating monochromatic x-radiation and simultaneous measurement by film and ionization chamber.

---

* A similar method adopted at about the same time in the Research Laboratories of Eastman Kodak Company, U.S.A., for sensitometry (private communication by H. E. Seemann and M. H. Horn) is interesting because, although known before the establishment of crystal reflection, the method has rarely been used.

directly opposite is fixed a smaller box with an aperture for admitting secondary radiation. The ionization chamber and the film are placed in the open side of the box opposite the aperture. The primary radiation impinges on a sheet of the secondary radiator (platinum, copper or a block of barium carbonate). The main purpose of the box is to prevent primary x rays from passing on to the film and to the ionization chamber, which were placed parallel to the plane of the aperture at a distance of 10 inches from the latter, as indicated in Figure 3. The characteristic radiations used for this investigation and the selective absorption filters to eliminate scatter and K_β radiations are given in Table 2.

Table 2

Wavelength (A.)	Radiation	Filter
0·188	K_α platinum	Selective filter $CaWO_4$.
0·39	K_α barium	Selective filter KI.
1·2	L platinum	No filter, excited by x rays generated at 70 kv., i.e. below excitation potential of K lines of platinum.
1·4	K_α copper	Selective absorption filter Ni.

These wavelengths were selected because they cover a fairly wide range and because the secondary radiators were easily obtainable.

The selective absorption filters were prepared as follows :

For the K_α platinum line a Kodak Ultra-Speed back intensifying screen was used. For the K_α barium line small crystals of potassium iodide were mixed with cellulose lacquer coated on card and placed between two sheets of x-ray film base. The weight of iodine per cm^2 required for the absorption of the K_β line was calculated on the basis of a transmission of 1/200 of the incident intensity. For the K_α copper radiation a nickel foil of 0·002 in. was used.

In order to check whether the radiations used were really monochromatic, a crystal spectrograph was designed for the purpose. The spectrograph was used with a stationary crystal as the crystal received radiation from different angles since the radiation from the secondary radiator was diffused. Spectrograms of the K lines of barium and platinum and the L lines of platinum show that the K_β line of barium was selectively absorbed by KI. In the case of platinum the K_α and K_β lines are not separately resolved, but it can be assumed that the K_β line was partly eliminated by the selective absorption of a sheet of calcium tungstate ($CaWO_4$). No background density due to scattered or continuous radiation is evident in the spectrograms.

A further check on the quality of the K radiation of barium and platinum was made by means of photographic absorption measurements using the method described by Seemann and MacGillivray (1946). The absorption curves obtained were almost linear. It was not possible to check whether the small deviation from linearity is due to a small background effect or to inaccuracy of the photographic method. The absorption coefficients of the two radiations were computed from the mean slopes of the absorption curves, and the equivalent wavelengths are set out in Table 3.

It is not possible to distinguish between the K_{α_1} and K_{α_2} lines, but the difference between them is in any case negligible. The wavelength used was therefore a mean of the K_{α_1} and K_{α_2} wavelengths weighted in the ratio 2 : 1 according to their relative intensities.

The quality of the platinum L radiation was checked by absorption measurements using stripped x-ray emulsions and also aluminium sheets of 0·1 mm. thickness. These absorption measurements were made by means of a 'grenz-ray' ionization chamber, which is provided with a 'Cellophane' window. The percentage absorption of one sheet of stripped emulsion derived from the absorption curve is 16% as against 17·8% when calculated from the known wavelength of platinum L radiation. The discrepancy between the two values may well be caused by the uncertainty in the mass absorption coefficient of silver bromide in this range of wavelengths. The mass absorption coefficient on which this calculation is based is 88·5.

In the case of the L radiation of platinum the effective wavelength was computed from a mean of the strongest L_α and L_β lines, from which a value of 1·2 A. was derived.

Table 3

Radiation	Wavelength (A.)	Wavelengths (A.) evaluated from photographic absorption curves
K_a platinum	0·178	0·183
K_a barium	0·39	0·37

Table 4. Monochromatic Radiation

Variation of density over film area

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A	0·47	0·54	0·58	0·61	0·62	0·61	0·61	0·60	0·58	0·55
B	0·57	0·61	0·64	0·66	0·65	0·65	0·64	0·63	0·61	0·60
C	0·58	0·63	0·66	0·68	0·70	0·69	0·68	0·67	0·65	0·65
D	0·59	0·66	0·69	0·73	0·74	0·74	0·72	0·71	0·68	0·67
E	0·61	0·66	0·72	0·75	0·78	0·79	0·77	0·75	0·73	0·71
F	0·58	0·64	0·70	0·73	0·76	0·77	0·76	0·75	0·72	0·71
G	0·59	0·63	0·68	0·73	0·76	0·77	0·75	0·75	0·71	0·68
H	0·57	0·63	0·68	0·72	0·75	0·76	0·75	0·73	0·69	0·65

No spectrogram was obtained of the CuK_α line owing to temporary difficulties with the spectrograph which had not been overcome when this paper was written. A check on the secondary radiation of Cu was made by means of ionometric absorption measurements, which showed that the radiation must have been slightly contaminated. It is probable therefore that the results on quantum efficiency using CuK_a radiation are less reliable than those obtained with the other radiations.

An essential test in regard to the uniformity of the diffused characteristic radiation thus obtained had to be carried out, since ionization chamber measurements and film strip exposures had to be made side by side simultaneously. The uniformity was examined by exposing to the monochromatic radiations a sheet of film which covered the positions of the plane of the film and the plane through the centre of the ionization chamber. The result obtained on the test film is shown in terms of densities measured at intervals over the film area of 6 in. × 8 in. in Table 4. The density values framed on the left of the Table cover the area in which the film strip was placed. The density values framed on the right represent the measures obtained in the plane covered by the centre of the ionization chamber. It will be seen that the mean of the densities obtained on the left and on the right agree fairly well, namely $D = 0{\cdot}68$ on the left and 0·72 on the right. As the centre of the chamber was placed flush with the plane of the film, the assumption was made that the dosemeter reading corresponds to the dose impinging on the film.

It may be further noted that the ionization measurements for the Pt_L and Cu K_α radiation, which are very soft, were made with the 'grenz-ray' chamber mentioned above. This was specially calibrated by the National Physical Laboratory for Pt_L radiation, using the same apparatus for producing the Pt_L radiation as is described above. The calibration was made against a free air standard ionization chamber.

§4. GRAIN COUNTING AND DENSITY MEASUREMENTS

The following types of emulsions were used for the evaluation of quantum efficiency: Type A, non-screen X-ray film; Type B, screen X-ray film; Type C, non-screen X-ray film of fine grain; Type D, chloro-bromide stripping emulsion. All films were processed with constant agitation in D19b X-ray developer at 68° F. for 4 minutes in a processing machine in which the reproducibility of development is of high order.

In the first instance, very small X-ray doses were given so that grain counts under the microscope could be made as described in §2. The number of grains per unit area was then plotted against the densities, which were read on a photoelectric densitometer of a high order of reproducibility. Typical plots which illustrate the relationship between the number of grains counted per unit area and the density are shown in Figure 4. The curves cut the ordinate at zero number of grains at a finite density corresponding to the density of the base. Although it was anticipated that the quality of radiation would have no influence on the relationship between the number of grains and density, a check was made by applying various qualities of radiation. The various points refer to radiations of different wavelengths, and are indicated by various symbols, as explained on the graphs of Figure 4. The random deviation of the points indicates that no influence of the quality of X rays could be observed. The regression lines relating the number of grains per unit area against density were evaluated by correlation methods. The standard deviation of points about regression lines is about 1 grain.

§5. EVALUATION AND RESULTS ON QUANTUM EFFICIENCY

The results of the measurements described above and plotted in Figure 4 were used for the evaluation of quantum efficiency. From the known value of the dose, measured in r-units for each exposure, the number of quanta absorbed per unit area in the various emulsions were computed, using equation (4) of §2, and the quantum efficiency was computed from the ratio:

$$\frac{\text{number of developed grains per unit area}}{\text{number of quanta absorbed per unit area}}.$$

Initially small X-ray doses below 0·01 r-units were given for the various wavelengths so that the number of developed grains could be counted. Since the doses were close to the limit of sensitivity of the dosemeter, greater accuracy was expected by using higher doses. In order to correlate higher dose measurements with the number of developed grains, the grain–density curves, Figure 4, were extrapolated to higher density values, and the number of developed grains was evaluated from density readings. The justification for the extrapolation of the grain number–density curve is given by the relationship derived by Nutting (1913),

$$D = 0{\cdot}4343\,\bar{a}n,$$

where D = density, $\bar{a}$ = the average projected area per grain and n = the number of developed grains.

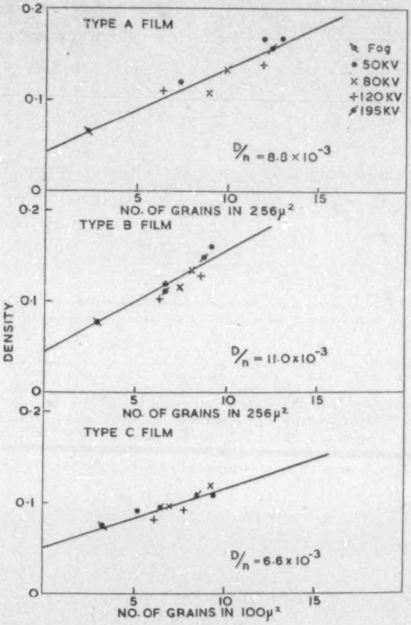


Figure 4. Relationship between photographic density and number of developed grains counted in Type A, B and C film for exposure to x rays generated at 50, 80, 120 and 195 kv.

Assuming that $\bar{a}$ remains constant as the exposure increases for densities below $D = 1$, the equation indicates proportionality between D and n. The constants of proportionality D/n of the straight line (Figure 4) were evaluated and are shown below the curves. The quantum efficiency was then computed from

$$\frac{n}{Q} = \frac{D}{r}\,\frac{n/D}{Q/r},$$

where n = number of developed grains per unit area above the fog, Q = number of quanta absorbed per unit area of single emulsion, r = number of roentgen units, and D = density above fog.

The ratios D/r and n/D were obtained experimentally as described above, and the ratio Q/r was computed from equation (4) § 2. The values of quantum efficiency obtained for the various frequencies of radiation and types of films are shown in the graphs (Figures 5–8) illustrating the grain yield per absorbed quantum η against the frequency ν of the radiation. Figure 5 shows this relationship for film of Types A, B and C for x rays over a range of wavelengths of

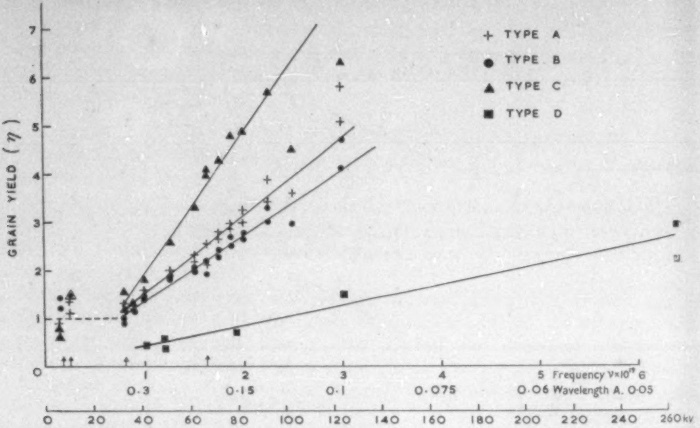


Figure 5. Relationship between grain yield and frequency of x rays for Type A, B, C and D film (exposures to monochromatic x rays are indicated on the abscissa axis by arrows).

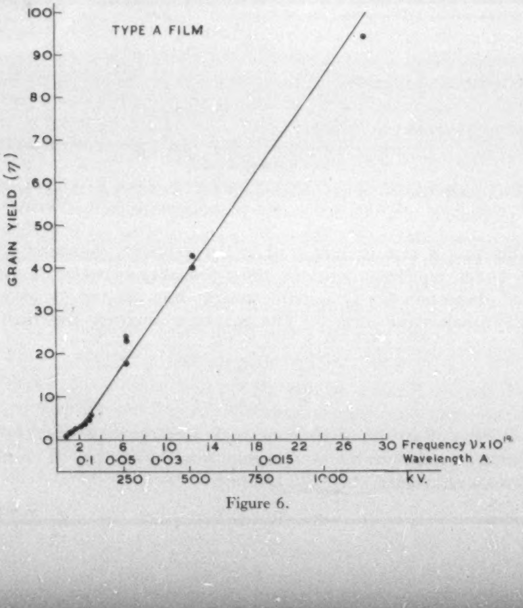


Figure 6.

1·54–0·1 A. and for Type D film over a range of 0·3–0·047 A. For Type D film the results have been obtained from actual grain counts, whereas the remaining results were derived from the extrapolated grain number–density curves (Figure 4).

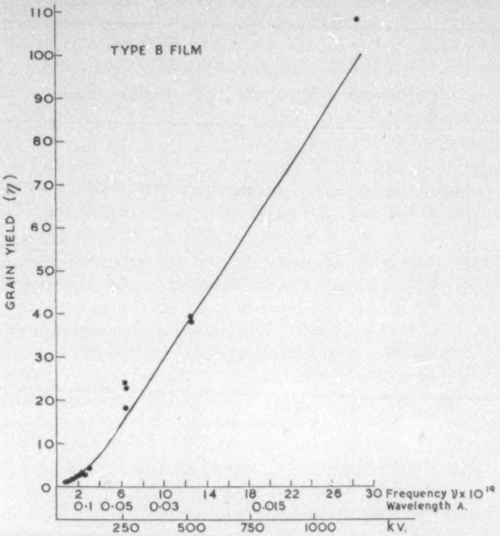


Figure 7.

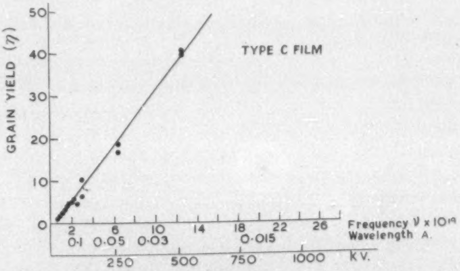


Figure 8.

Figures 6–8. Relationship between grain yield and frequency of x rays generated between 40 and 1,250 kv. (Type A, B and C film).

In order to cover also the results obtained with very hard radiation such as is generated at 475 kv. and 1,000 kv. and with gamma rays from a radium source, graphs are presented in Figures 6, 7 and 8, Types A, B and C film respectively.

It should be noted that the mean wavelength corresponding to gamma rays emitted from a radium source filtered by 1 cm. of lead is quoted as 0·0105 A., which is equivalent to 1,230 kv.

An interpretation of the results shown in Figures 5–8 is dealt with briefly in the following section and in more detail in §7.

§6. QUANTUM EFFICIENCY IN WATER-SWOLLEN FILMS AND CONSIDERATIONS OF SUB-IMAGE FORMATION

(i) *Quantum Efficiency in Water-Swollen Films*

The experimental result that several grains may be rendered developable per quantum absorbed can be understood on the assumption that the energy of the absorbed quantum is sufficient to release photoelectrons from the atoms of the affected grain which, during their passage through the emulsion, render other neighbouring grains developable. Some evidence for this effect was given by the exposure of Type A film in the wet state, where it can be assumed that the grains are further apart than in the dry state. In the wet state photoelectrons emitted from an affected grain are more readily absorbed by the water-swollen gelatin, and will be less able to reach neighbouring grains, also because of the greater average separation between grains. The quantum efficiency found for the film exposed in the wet state at a wavelength of 0·1 A. was found to be almost half of that found if the same emulsion was exposed in the dry state. This result seems to indicate that the decrease of grain yield is caused by the reduced length of path of photoelectrons, the water-swollen emulsion and the greater separation between grains (Herz 1949 a).

In the region of grain yield 'unity' it may be assumed that the path length of photoelectrons released from the affected grain is not sufficient to reach neighbouring grains, or at least that these photoelectrons do not contribute to the developability of neighbouring grains. Evidence for this assumption was given by exposure of Type A film in the wet state at $\lambda = 0{\cdot}37$ A., where the grain yield is unity in the dry state of the film. The result showed that the same grain yield of unity was obtained in the wet state as in the dry state. It can be thus concluded that a grain yield of unity is obtained by single quantum hits without contributions of photoelectrons from neighbouring grains.

(ii) *Considerations of Sub-Image Formation*

It was shown by Burton and Berg (1946) that photographic emulsions showing a toe in a density–exposure curve are likely to show sub-image formation. The only emulsion used in these tests which showed a slight indication of a toe in the density–exposure curve was Type D film, which also proved to have a quantum efficiency below unity for wavelengths longer than 0·13 A. A low-intensity post-exposure of this film to light gave evidence for sub-image formation which could not be found in any of the other emulsions.

§7. INTERPRETATION AND CONCLUSIONS

From a study of Figures 5–8 the following conclusions can be drawn:

(i) *Grain yield between 1·2–0·37 A.* (Figure 5). In Figure 5, a flat part of the curves at a quantum efficiency of 1 is seen in the region of very soft radiation corresponding to a wavelength range between 1·54–0·37 A. The energy of one absorbed quantum is evidently sufficient to render one grain developable over this relatively wide range of wavelengths in each case of the three emulsions Types A, B and C. This means that the number of ionizations produced from one quantum absorbed occurring at the shorter wavelength 0·37 A. is more than is required to render one grain developable. The range of photoelectrons, however, appears to be insufficient for them to reach neighbouring grains. Thus, no contribution to the developability of neighbouring grains is made by one quantum hit. This supposition is further supported by the experimental result described in the last section, which showed that the quantum efficiency of the water-swollen emulsion at 0·37 A. remains at unity.

(ii) *Grain yield between 0·37–0·0105 A.* (Figures 6–8). The most characteristic feature of the graphs appears to be that the grain yield is proportional to the quantum energy within wavelengths from 0·37–0·1 A. and for the Type C film down to a wavelength of 0·047 A. The curves (Figures 6–8) do not pass through the origin, and those for films Types A and B bend at a wavelength about 0·1 A.

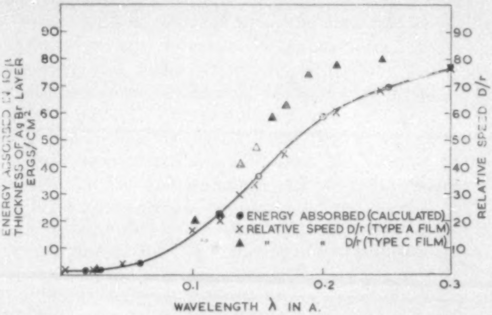


Figure 9. Relationship between energy absorbed (erg/cm²) in 10 μ layer of silver bromide and X-ray wavelength, and between speeds of Type A and C film and wavelength.

The results agree with those obtained by Pelc (1945). His grain yield figures are however higher than those reported in this paper by a factor of 8, due to an erroneous silver mass determination (private communication from Pelc).

In order to compare this result with the theoretical expectation (§ 2) reference may be made to the Figures 9 and 10. Figure 9 shows the relationship between wavelength and the values of the energy absorbed per unit area in a 10 μ layer of silver bromide (as calculated from equation (2) § 2). Figure 10 shows the calculated values of the number of quanta absorbed per unit area in a 10 μ layer of silver bromide (using equation (4) § 2) as function of the wavelengths. The experimental relative speeds for films Types A and C are expressed in values of D/r, where D is the density and r the number of roentgen units, and were adjusted to fit the value of the energy absorbed at a wavelength of 0·3 A. (Figure 9) and to fit the number of quanta absorbed at a wavelength of 0·3 A. (Figure 10). The speed values refer only to measurements taken within the straight line part of the density–exposure curve of Type A and C films.

By reference to Figure 9 it is seen that the speed values for Type A film coincide with the values of the energy absorbed over the complete range of wavelengths.

Since the quantum energy rises with shorter wavelength, the number of grains rendered developable per absorbed quantum must evidently be proportional to the quantum energy absorbed at each wavelength. It will be noted however from the curve of Type C film (Figure 9) that the relative speed values of this film do not agree so closely with the values of the absorbed energy over the range of wavelengths considered.

In Figure 10 it is seen that the deviation of the speed values from the number of quanta absorbed increases with shorter wavelengths. This indicates that more grains are rendered developable per absorbed quantum with short than with long wavelengths.

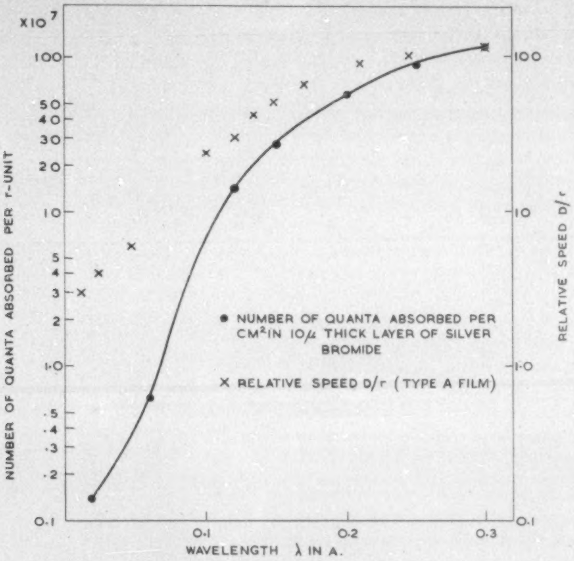


Figure 10. Relationships between the number of quanta absorbed per r-unit, relative speed of Type A film and x-ray wavelength.

The occurrence of the bend in the curves in Figures 6 and 7 at a wavelength of about 0·1 A. is not understood. It may be a real effect, as it does not occur for Type C film. It is not likely to be an observational error since the same method and data were also used for the evaluation of Type C film (Figure 8).

The mechanism of the increase of the number of grains rendered developable with increasing quantum energy can be interpreted as follows: (*a*) With increasing quantum energy photoelectrons of increasing energy are released. The energy of the emerging electrons will be $h\nu - w_K$, where w_K is the energy required to remove an electron from the K level. (*b*) The range of photoelectrons thus released by a single quantum hit increases with quantum energy, thus affecting an increasing number of grains adjacent to the quantum absorbing grains.

The mean energies of photoelectrons plus recoil electrons released in a silver bromide emulsion for various wavelengths of incident radiations have been computed, and the corresponding ranges of electrons in gelatin and in silver bromide emulsions are known approximately. Experimental values of the latter were found in recent work on the range of electrons in nuclear track emulsions (Herz 1948, 1949 b). At the critical wavelength of 0·37 A. (at a grain yield of unity), where the energy of photoelectrons appears to be insufficient to affect neighbouring grains, the mean electron energy has been found to be about 9 kv. A 9 kv. electron has a range of about 0·9 μ in a silver bromide emulsion of the Type A. It would thus lose its energy entirely before it reaches a neighbouring grain, as the spacings of the grains in Type A emulsion were found to be of the order of 1–2 microns. At a quantum efficiency of 2, corresponding to a wavelength of 0·23 A. for Type A film, the mean electron energy was found to be 32 kv., corresponding to a range of 9 μ in the silver bromide emulsion. This relatively large range appears evidently to be sufficient for the electrons to contribute to the developability of adjacent grains.

Summarizing, it may be said that about the same fraction of the energy is being utilized for latent image formation at all points in the range from low to high quantum energies. It is further interesting to note that as a consequence of this as many as 80 grains at 1,000 kv., and 100 grains in the range of gamma rays are rendered developable per one quantum absorbed.

(iii) *Rate of increase of grain yield with quantum energy*. The results of the grain yield of three different emulsions (Figure 5) are indicated by the different slopes of the curves. The rate of increase of grain yield with quantum energy is considerably steeper for Type C film than for Type A and B films, the slopes of which are fairly similar. It is believed that this higher rate of increase for Type C film is caused by the closer packing of grains in this emulsion. Considering the closer packing of grains in Type C film, it seems evident that for photoelectrons of the same range released from one quantum-absorbing grain, more grains may be affected than in an emulsion with looser packing of grains.

If a certain quantum energy (at about 0·05 A.) is reached, the slopes of the curves (Figures 6–8) become practically identical for the three emulsions. This probably means that if a certain length of path of photoelectrons is reached, the difference in spacing of the grains loses its importance in regard to grain yield, and similar yields are obtained for all three emulsions.

(iv) *Grain yield for type D film* (*Figure* 5). In Figure 5 it is shown that Type D film reaches a grain yield of unity only at relatively short wavelengths and is well below unity for wavelengths longer than 0·12 A. This implies that the grains in Type D film are relatively insensitive. It appears that one quantum hit is not sufficient to render one grain developable, and that many grains, although affected, do not receive a sufficient number of ionizations to render them developable. Thus sub-images are formed, evidence of which was given by low intensity light post-exposure in § 6.

By comparing the results of quantum efficiency of Type D film in the region of grain yield below unity, at a wavelength of 0·12 A., with those of films of Types A, B and C (Figure 5), it may be at first surprising that no flat region in the grain yield–frequency curve is obtained for this film. This may be understood by the following considerations. At wavelengths longer than 0·12 A. the range of the photoelectrons released is sufficient for them to reach grains adjacent to the quantum-absorbing grain, as is shown by reference to the results obtained with other films. The energy of these photoelectrons is however insufficient to render the grains of Type D film developable (sub-image formation). At wavelengths shorter than 0·125 A. it appears that not only the range but also the energy of the photoelectrons becomes great enough to produce sufficient ionizations and thus render neighbouring grains developable.

REFERENCES

Burton, P. C., and Berg, W. F., 1946, *Photogr. J.*, **86B,** 1–24.
Eggert, J., and Noddack, W., 1927, *Z. Phys.*, **23,** 222.
Günther, P., and Tittel, H., 1933, *Z. Elektrochem.*, **39,** 646.
Herz, R. H., 1948, *Nature, Lond.*, **161,** 928; 1949 a, *Photogr. J.*, **898,** 89; 1949 b, *Phys. Rev.*, **75,** 478.
Mayneord, W. V., 1940, *Brit. J. Radiolog.*, **13,** 235.
Nutting, P. G., 1913, *Phil. Mag.*, **26,** 423.
Pelc, S. R., 1945, *Proc. Phys. Soc.*, **57,** 523.
Seemann, H. E., and MacGillivray, L. L., 1946, *Rev. Sci. Instrum.*, **17,** 12.
Silberstein, L., and Trivelli, P. H., 1930, *Phil. Mag.*, **9,** 787.

The Analysis of Observations on Spaced Receivers of the Fading of Radio Signals

By B. H. BRIGGS, G. J. PHILLIPS and D. H. SHINN
Cavendish Laboratory, Cambridge

Communicated by J. A. Ratcliffe; MS. received 2nd August 1949

ABSTRACT. The fading of a radio wave once reflected from an irregular ionosphere is discussed in terms of the variable diffraction pattern produced at the ground. It is pointed out that fading may arise either by a drift of the diffraction pattern past a receiver, or by irregular variations in the pattern, or by both mechanisms together. It is shown how, from observations at three receiving points, it is possible to deduce the rate at which the pattern is changing and the velocity with which it drifts over the ground. This velocity is of interest as it may be related to the wind velocity at ionospheric heights in the atmosphere, but the relation of the diffraction pattern to the irregular ionosphere which produces it is not discussed in detail.

§ 1. INTRODUCTION

Several papers (Ratcliffe and Pawsey 1933, Pawsey 1935, and more recently Ratcliffe 1948 and Mitra 1949) have drawn attention to the fluctuations in the amplitude of a radio wave which has been singly reflected from the ionosphere. It has been pointed out that the reflection of a radio wave from an irregular ionosphere will lead to the production of an irregular 'diffraction pattern' on the ground. For practical reasons this pattern can only be observed at a few fixed points, by means of spaced receivers. If the ionosphere remained constant and fixed in position, the amplitude of the wave at each point on the ground would remain constant, and there would be no fading.

One way in which fading could be produced would be by a movement of the diffraction pattern past the observing points, as would happen if there were a horizontal wind in the ionosphere. In this case the fading recorded at two points separated in the direction of the wind would be exactly similar but displaced in time, and from the time difference the velocity of the wind could be deduced.* If, now, the receivers were also gradually separated in the direction perpendicular to the direction of the wind, the recorded fading would become less similar, until finally no recognizable similarity would remain.

Another way in which fading could be produced would be by *random* changes in the ionosphere, as, for example, by the random motion of different parts of the ionosphere (Ratcliffe 1948). Such a mechanism would produce random changes of the diffraction pattern on the ground. The fading at two nearby receivers would then be similar, but would become more dissimilar as the receivers were separated from each other in any direction. There would be no tendency for the occurrence of similar fading with a time lag.

In general the two mechanisms mentioned above may operate simultaneously, so that, although the diffraction pattern moves on the ground with a steady drift velocity, it is subject to random changes of form as it moves. Under these circumstances the fading recorded by two receivers separated in the direction of the wind would be similar but displaced in time if the distance of separation were not too great, but the degree of similarity would decrease as the separation was increased, until at sufficient distance all similarity would have disappeared.

In this paper we shall show that, if certain plausible assumptions are made, it is possible to deduce the following from fading records taken at three points situated at the corners of a right-angled triangle: (*a*) the steady velocity with which the diffraction pattern drifts over the ground; (*b*) the rate at which the pattern alters as it moves; (*c*) the size of the irregularities in the pattern. The principle of the method is to consider the general similarity of the three records, and from the extent to which the differences exceed those expected from a drifting pattern of fixed character one can find the rate at which the pattern alters; hence one can make the correction that is necessary to the value of drift as it appears from the time lag. We need to consider the auto-correlation function of the fading records at the three points considered and also the cross-correlation functions between pairs of these records, and we devote the next section (§2) to a definition of these functions and a short discussion of them. The first step in solving our problem is to prove relations between certain fundamental quantities defined in §4, and this is done by two different methods which lead to the same results. Thus in §5 we discuss the correlations between the fading at the three receivers in a direct manner, focusing attention on the way in which the amplitude of signal varies with position and time, while in §6 we consider the form of a 'correlation surface' which shows how the correlation between the amplitudes at two receivers varies with their separation and with the difference in times of observation. It then remains to show in §7 how the results of the previous sections can be used to obtain the fundamental characteristics of the pattern from the three fading records.

The results which we deduce relate entirely to the diffraction pattern; in particular the steady drift and the random changes which we discuss are those of the pattern on the ground. We have deferred the interesting problem of relating this pattern to the phenomena in the ionosphere that give rise to it. By so doing we avoid having to base the discussion on any particular model of the ionosphere, although at certain stages in the analysis we do make some simplifying assumptions about the nature of the diffraction pattern. It is sufficient to remark that drifting effects in the diffraction pattern are probably connected with large scale movements in the ionosphere and changes in the pattern with random movements in the ionosphere.

---

* If a simple diffraction mechanism is operative, the wind in the ionosphere will be half the velocity of drift of the diffraction pattern over the ground (Pawsey 1935).

Although we naturally consider the amplitude of a radio signal as the basic quantity when discussing fading, the results could be applied to any other suitably varying characteristic of an ionospheric reflection. Munro (1948) has shown how the magnitude of the magneto-ionic splitting, for example, is useful in the investigation of movements in the F region.

Some practical results concerning the drift of patterns over the ground have already been obtained (Mitra 1949). These experiments are being continued, with modified apparatus, and we hope to publish the results in the near future.

§ 2. MEASUREMENT OF CORRELATION

(i) *A Single Time-Record*

A record of fading, i.e. a graph of the amplitude R of a reflected wave against time t, is found to have a random character like the curve in Figure 1 (a). An important function which arises in the study of a variable of this kind is the *auto-correlation function.* Thus, corresponding to the original random function $R(t)$, we have the auto-correlation function:

$$\rho(\tau)=\frac{\overline{\{R(t)-\overline{R}\}\{R(t+\tau)-\overline{R}\}}}{\overline{\{R(t)-\overline{R}\}^2}}. \qquad \ldots\ldots (1)$$

This function measures the average correlation between values of R separated by a time interval τ. It has the value unity for $\tau=0$, and in all practical cases decreases smoothly at first as τ increases from zero. A plot of a typical auto-correlation function (or auto-correlogram) is shown in Figure 1 (c). The value of τ at which $\rho(\tau)$ falls to zero is clearly that time separation at which two values of R have become completely independent.

In the analysis of fading of radio signals we need a suitable measure of the 'speed of fading'. It is possible to derive such a measure from the function ρ just defined, but it is more convenient to define the speed of fading, S, as follows:

$$S=\left|\overline{\frac{\partial R}{\partial t}}\right| \Big/ \overline{R}. \qquad \ldots\ldots (2)$$

This is simply the mean rate of change of R, normalized to a constant mean value of R.

(ii) *Space Variations, and the Two-Dimensional Auto-Correlation Function*

By analogy with the definition of S we may measure the rate of variation of R over the ground at a given instant by a quantity G, where

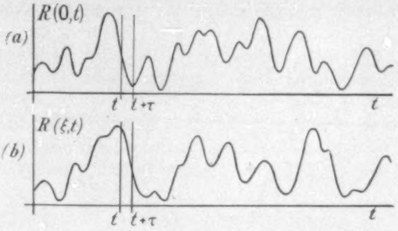
$$G=\left|\overline{\frac{\partial R}{\partial x}}\right| \Big/ \overline{R}. \qquad \ldots\ldots (3)$$

We shall assume that G is independent of the direction over the ground in which the variable x is measured. (An extension of the theory in which this condition is removed is mentioned in § 8.)

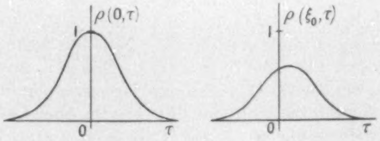
In the practical case the signal amplitude over the ground and in time will be represented by a function $R(x, y, t)$ of the two space coordinates x, y, and the time coordinate t. In order to simplify the analysis in the first instance, we shall introduce the artificial idea of a *one-dimensional* ground, so that R is a function of two variables only, x and t. We can derive the two-dimensional auto-correlation function of $R(x, t)$ by an obvious extension of equation (1). Thus

$$\rho(\xi, \tau) = \frac{\overline{\{R(x, t) - \overline{R}\}\{R(x + \xi, t + \tau) - \overline{R}\}}}{\overline{\{R(x, t) - \overline{R}\}^2}}. \qquad \ldots\ldots (4)$$

The example of a one-dimensional ground is useful for purposes of illustration because the functions $R(x, t)$ and $\rho(\xi, \tau)$ can be visualized as surfaces.



(a) First Receiver (b) Second Receiver



(c) Auto-correlation (d) Cross-correlation

Figure 1.

In a hypothetical experiment on our one-dimensional ground we might obtain time records of the amplitude as shown in Figure 1 (*a*) and (*b*), at two fixed points, say at $x = 0$ and $x = \xi_0$. We are then limited to a knowledge of $R(0, t)$ and $R(\xi_0, t)$. From either record the auto-correlogram corresponding to $\rho(0, \tau)$ and shown in Figure 1 (*c*) may be found. In addition, it is useful to evaluate the cross-correlation between the two records for different values of the time lag τ, so as to obtain the *lag-correlogram* corresponding to $\rho(\xi_0, \tau)$ shown in Figure 1 (*d*). These curves may be thought of as sections of the ρ surface by planes $\xi = 0$ and $\xi = \xi_0$. Since in the practical case we know the amplitude variations in time much more completely than those along the ground, we must regard the averaging in obtaining ρ from equation (4) as being carried out over time only. The physical problem is such that it may reasonably be assumed that the same result will be obtained whether we average over space and time, or over time only. This assumes that we possess statistically stationary records of sufficient length to include many cycles of the variation of R.

The extension of these ideas to cover a two-dimensional ground is straightforward. We then consider a correlation function $\rho(\xi, \eta, \tau)$ derived from $R(x, y, t)$ and $R(x+\xi, y+\eta, t+\tau)$.

§ 3. STATEMENT OF THE PROBLEM

Before proceeding with the mathematical discussion, and in order to introduce some of the difficulties encountered in the general case, we consider a two-dimensional diffraction pattern which is not subject to random changes. Consider the uniform drift of such a constant ground pattern past two observing points fairly close to each other. In Figure 2, let P and Q be two such points of observation on a line OX, and let the form of the drifting pattern be represented by contours of constant amplitude. We assume that, on the average, these contours are circles, i.e. the pattern changes equally rapidly in all directions. Let the pattern move in a direction making an angle ϕ with OX. Now the detail which passes P can never recur in exactly the same form at Q (unless ϕ is zero) but a somewhat similar variation is obtained at Q, with a time lag corresponding to the time taken for the pattern to traverse the projection of PQ on the line of motion of the pattern. As a result of the measurements at P and Q, then, we observe an *apparent* velocity V'_{OX} along OX. This quantity V'_{OX} is defined as the length PQ divided by the observed time lag, and is easily shown to be given by

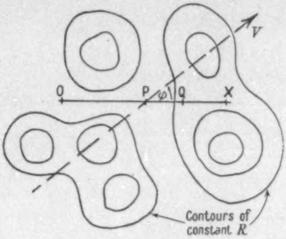


Figure 2.

$$V'_{\mathrm{OX}} = V/\cos\phi. \qquad \ldots\ldots (5)$$

It will be noted that this apparent velocity along OX is greater than the true veloctiy V of the pattern.

We shall find, however, that the velocity V_{OX} with which an observer would have to move along OX in order to reduce the speed of fading as observed by him to a minimum is given by

$$V_{\mathrm{OX}} = V\cos\phi.$$

The difference between V'_{OX} and V_{OX} increases as ϕ increases.

The important conclusion to be drawn is that we have two methods of interpreting the observations along OX as a drift. The latter method is the one which gives the true component of the total drift velocity along OX. But it may be more convenient in practice to make measurements leading to the apparent velocity V'_{OX} of the first method. If this is the case we must remember that we have found $V/\cos\phi$ and not $V\cos\phi$.

In the following sections we shall seek relations between the various velocities, like V'_{OX} and V_{OX} above, which arise in the more general case of a moving pattern with random changes, and consider how they are obtainable from records made on spaced receivers. It is possible to obtain the relations in more than one way. We shall therefore pass on immediately in the next section to the definitions of the velocities and other parameters which seem fundamental. Having introduced these, we shall in §5 derive several useful formulae using a modification of the $R(x, t)$ surface which was mentioned in the discussion of the definition of $\rho(\xi, \tau)$. This approach may be called that of the $(x, V_c t)$ diagram. Then in §6 a shorter alternative approach will be given which is somewhat more mathematical. Here the formulae are proved from the geometry of the $\rho(\xi, \tau)$ surface. In both these approaches we prove the results first for the 'one-dimensional ground' introduced in the previous section, and then generalize to the practical case of a two-dimensional ground.

We shall assume that the contours of constant ρ in the $\rho(\xi, \tau)$ diagram are similar ellipses concentric about the origin. The significance of this assumption is discussed later in §8. Similarly when we consider a two-dimensional ground we shall take the constant ρ surfaces in the $\rho(\xi, \eta, \tau)$ diagram as similar concentric ellipsoids, which intersect the ξ–η plane in circles. In this way we include the assumption of an isotropic ground pattern.

§4. DEFINITIONS

The following definitions of various basic quantities are given in words so that they may apply generally in the cases of a one- or two-dimensional ground without being associated with any particular mathematical approach. In the illustration of the definitions, however, use is made of the correlation function for a one-dimensional ground, and it may be helpful to refer to Figure 4 (*c*) which is a plan of the $\rho(\xi, \tau)$ surface showing the contour $\rho = \rho_0$, and to Figure 4 (*b*) showing vertical sections corresponding to the correlograms of Figures 1 (*c*) and 1 (*d*).

(i) *Fading velocity* V_c'. This is a measure of the ratio of space shift to time shift needed to produce, on the average, the same change in the value of R. Thus

$$V_c' = x_0/t_0, \qquad \ldots\ldots (6)$$

where x_0, t_0 satisfy $\rho(x_0, 0) = \rho(0, t_0)$.

Under the simplifying assumptions mentioned above this definition does not depend on the value of ρ chosen, or on the direction over the ground in which the space coordinate denoted by x is measured. Also, in view of (2) and (3), the definition leads to the expression

$$V_c' = S/G, \quad \text{or} \quad S = GV_c'. \qquad \ldots\ldots (7)$$

We thus get the physical interpretation that the fading velocity V_c' is the velocity with which the unchanging ground pattern would have to move past an observer to give the observed speed of fading, i.e. the velocity of drift that would be needed to explain the fading *entirely* in terms of a drifting pattern with no random changes.

(ii) *Drift velocity V.* This is the velocity of an observer who has so adjusted his motion over the ground that he experiences the slowest possible speed of fading. Suppose he compares signal amplitudes at times τ_1 apart. His displacement ξ_1 during this time must be adjusted until $\rho(\xi_1, \tau_1)$ is a maximum, and then

$$V = \xi_1/\tau_1. \qquad \ldots\ldots (8)$$

This velocity is an unambiguous measure of the movement of the pattern along the ground. For a two-dimensional ground V has direction. Assuming axes of x and y on the ground let the direction be at angle ϕ to the x axis and let components of the velocity along the axes be V_x and V_y.

(iii) *Characteristic velocity, V_c, and characteristic speed of fading S_c.* These are the values of V_c' and S which are found by the observer above moving at velocity V, and give a quantitative measure of the amount of random change taking place. To this observer the ratio of space shift and time shift needed to produce a similar change in R is

$$V_c = x_0/\tau_1, \qquad \ldots\ldots (9)$$

where $\rho(x_0, 0) = \rho(V\tau_1, \tau_1) = \rho(\xi_1, \tau_1)$. This is seen by noting that any time shift τ must be accompanied by a displacement $V\tau$ for this observer. Also we have the relation corresponding to (7)

$$V_c = S_c/G, \quad \text{or} \quad S_c = GV_c. \qquad \ldots\ldots (10)$$

(iv) *Apparent drift velocity, V', and time lag τ_0 for best cross-correlation.* If the fading records of receivers separated by distance ξ_0 are examined for cross-correlation, i.e. if $\rho(\xi_0, \tau)$ is found for various values of τ, there will be a maximum correlation for a certain delay τ_0. The apparent velocity of drift is then

$$V' = \xi_0/\tau_0. \qquad \ldots\ldots (11)$$

This velocity is analogous to the velocity V'_{0x} discussed in § 3, and is the value that at first sight appears to be the true velocity of motion of the pattern, assuming this to be along the line of the receivers. For two-dimensional ground, we suppose that we have receivers spaced ξ_0 along Ox and spaced η_0 along Oy, and that τ_{0x} and τ_{0y} are the respective lags for best correlation. Then the apparent components are given by

$$V'_x = \xi_0/\tau_{0x} \quad \text{and} \quad V'_y = \eta_0/\tau_{0y}. \qquad \ldots\ldots (12)$$

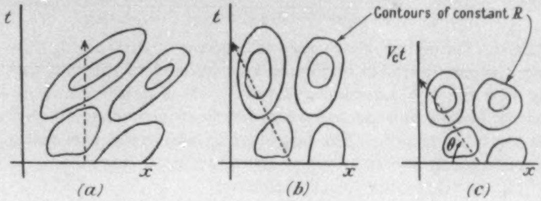
It is important to realize that these are not components of a vector V', but two quantities like V' defined from observational data.

§ 5. DEDUCTIONS FROM THE $(x, V_c t)$ DIAGRAM

(i) *Construction of the Diagram*

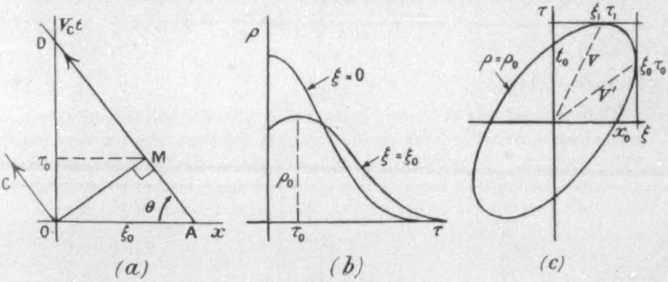
We imagine plotted, by means of contours of R, the complete account of the happenings over a one-dimensional ground as time progresses. In general we obtain something like Figure 3 (*a*). There is a tendency for the contours to be parallel to a line whose slope depends on the velocity of drift of the maxima and minima of R over the ground. If the origin of the x coordinate is moved along with velocity V we obtain Figure 3 (*b*), corresponding to the state of affairs seen by the observer moving with velocity V, who experiences the minimum speed of fading. We have seen that for such an observer the ratio of the space and time shifts which have equal effect (on the average) in changing R is V_c. Therefore we produce the third diagram in Figure 3 (*c*), using the coordinate $V_c t$ in place of t so that the average gradient of the contours is the same along both axes of the diagram.

Now we assume that in this diagram the average change, or expected correlation, between values of R at any two representative points depends only on their distance apart on the diagram. In other words a correlation surface corresponding to this new diagram would have contours in the form of circles centred on the origin. It can be seen by the reversal of the transformation back to Figure 3 (*a*), that on this assumption the real $\rho(\xi, \tau)$ contours must be ellipses.



(*a*) (*b*) (*c*)

Figure 3.
Dotted lines cover the values of R experienced by a fixed point.



(*a*) (*b*) (*c*)

Figure 4.

A fixed point on the ground will have velocity $-V$ relative to the axes of Figure 3 (*c*). Its representative point on the diagram must move towards the negative x direction as time t increases; in fact it traverses a line sloping backwards at an angle θ where

$$\cot\theta = x/V_c t = V/V_c. \qquad \ldots\ldots (13)$$

Thus the values of R along this line are those experienced at the fixed point on the ground. The $(x, V_c t)$ diagram is redrawn in Figure 4 (*a*), in which the contours have been omitted for clarity.

(ii) *Expression for the Speed of Fading*

Since for a given time interval the distance along OC is greater by a factor $1/\sin\theta$ than along the $V_c t$ axis, we have a comparison of the corresponding speeds of fading S and S_c. Thus

$$S = S_c/\sin\theta = S_c[1 + (V/V_c)^2]^{\frac{1}{2}}$$

using (13) to eliminate θ. If we consider V_c' and V_c as measures of the speed of fading we have in view of (7) and (10)

$$V_c' = V_c[1+(V/V_c)^2]^{\frac{1}{2}}.$$

The last two equations may be re-written

$$S^2 = S_c^2 + G^2V^2, \quad \text{and} \quad V_c'^2 = V_c^2 + V^2. \qquad \ldots\ldots (14)$$

The expression for the resultant speed of fading thus involves a square law combination of the contributions S_c and GV due respectively to random changes and steady drift.

(iii) *Time Lag for Optimum Cross-Correlation*

We obtain the fading record of a second receiver at distance ξ_0 from the first by following a sloping line AD in Figure 4 (*a*) parallel to the first line OC and intercepting the x axis at A, a distance ξ_0 from O. It is seen that a value of R at O on the first record would be expected to be most nearly correlated with the closest point M on the other record. The height of M above the axis will give $V_c\tau_0$, where τ_0 is the time lag corresponding to optimum cross-correlation. From the geometry of Figure 4 (*a*) we arrive at the equation

$$V_c\tau_0 = \xi_0 \sin\theta \cos\theta = \xi_0 V_c V/(V_c^2 + V^2)$$

in which we again make use of (13) to eliminate θ. Thus

$$V' = \xi_0/\tau_0 = (V_c^2 + V^2)/V$$

or

$$V'V = V_c^2 + V^2,$$

and hence from (14)

$$V'V = V_c'^2. \qquad \ldots\ldots (15)$$

It will be noticed that the determination of τ_0 gives V' immediately, but that V_c' is required to find V. This can be found as follows: the cross-correlation $\rho(\xi_0, 0)$ between the two records gives a measure of the effect of a pure space shift ξ_0, and if we find the time lag τ_s which gives the same value of the correlation coefficient $\rho(0, \tau_s)$ measured on one record, then the definition of V_c' gives

$$V_c' = \xi_0/\tau_s. \qquad \ldots\ldots (16)$$

(iv) *Other Methods*

It is useful to consider the time τ_e for which the auto- and cross-correlograms cross one another when plotted on the same axes as in Figure 5 (*b*). Mathematically τ_e may be defined as satisfying

$$\rho(0, \tau_e) = \rho(\xi_0, \tau_e) = \rho_e.$$

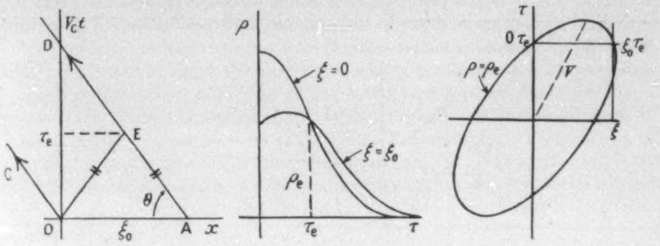
If we imagine the function $\rho(\xi, \tau)$ to be represented by a surface, then Figure 5 (*b*) represents the section by planes $\xi = 0$ and $\xi = \xi_0$. The contour $\rho = \rho_e$ of the degree of correlation in this case is shown in Figure 5 (*c*), which is a plan of the surface looking down on the (ξ, τ) plane.

The interpretation of τ_e in the $(x, V_c t)$ diagram of Figure 5 (*a*) is that it corresponds to the point E on AD such that OE = AE, since with this value of time lag the auto-correlation corresponding to AE equals the cross-correlation corresponding to OE. By simple geometry D is twice the height of E above the x axis so that OD is $2V_c\tau_e$. Thus we have $\cot\theta = \xi_0/2V_c\tau_e$. But from (13) $\cot\theta = V/V_c$. Hence

$$V = \xi_0/2\tau_e. \qquad \ldots\ldots (17)$$

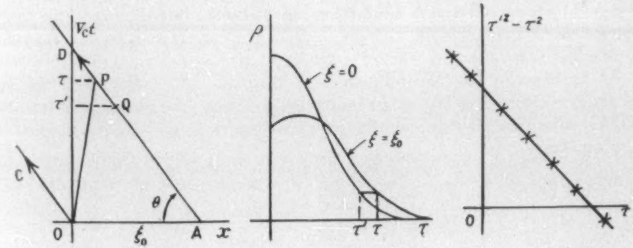
If the complete correlation curves can be worked out and plotted, the parameters V and V_c' are best obtained by the following method (Figures 6 (*a*) and (*b*)). Here a general point P on the record AD is considered with time lag τ relative to O or A. The auto-correlation for lag τ is thus governed by distance AP, and the cross-correlation for lag τ by distance OP. We have $AP = V_c\tau/\sin\theta$. Also let Q be a point on the line AD corresponding to time τ', so that AQ gives the auto-correlation for lag τ'. Now make AQ = OP, i.e. let us define τ' as the time lag giving an auto-correlation equal to the cross-correlation for lag τ. Then we also have

$$OP = AQ = V_c\tau'/\sin\theta.$$


V_ct
D
τ_e
E
C
θ
O
ξ_0
A
x
ρ
$\xi=0$
$\xi=\xi_0$
ρ_e
τ_e
τ
τ
$0\ \tau_e$
$\xi_0\tau_e$
$\rho=\rho_e$
V
ξ

(*a*) (*b*) (*c*)

Figure 5.


V_ct
D
τ
P
τ'
Q
C
θ
O
ξ_0
A
x
ρ
$\xi=0$
$\xi=\xi_0$
τ'
τ
τ
$\tau'^2-\tau^2$
0
τ

(*a*) (*b*) (*c*)

Figure 6.

From equations (13) and (14) $V_c/\sin\theta = V_c'$, and in triangle OPA of Figure 6 (*a*)

$$OP^2 = AP^2 + OA^2 - 2OA\,.\,AP\cos\theta.$$

Combination of the last four results yields

$$V_c'^2\tau'^2 = V_c'^2\tau^2 + \xi_0^2 - 2\xi_0\cos\theta\,.\,V_c\tau/\sin\theta.$$

Hence, using (13)

$$\tau'^2 - \tau^2 = (\xi_0^2 - 2V\xi_0\tau)/V_c'^2. \qquad \ldots\ldots (18)$$

This shows how use may be made of the complete auto- and cross-correlation diagrams to plot a straight line graph as shown in Figure 6(*c*), in which $\tau'^2 - \tau^2$ is plotted against τ. The intercepts which the line makes on the two axes are readily shown to be equal to $\xi_0/2V = \tau_e$ and $\xi_0^2/V_c'^2 = \tau_s^2$. Again V and V_c' may be obtained fairly simply from ξ_0. The region where a good straight line is obtained is an indication of the region in which the assumptions of the present treatment are valid.

(v) *Extension for Two-Dimensional Ground*

Some of the above equations need modification before they will apply for a two-dimensional ground, unless the two receivers happen to be in line with the drift velocity. To obtain more general formulae we consider the natural extension of the $(x, V_c t)$ diagram to the $(x, y, V_c t)$ diagram shown in Figure 7. The fading records of three receivers at A, O and B are given by the intersection of the three sloping lines OC, AD and BE with the surfaces of constant R in $(x, y, V_c t)$ space. It is assumed that variations of R are equal on the average for equal displacements in any direction in this three-dimensional diagram. Considerations similar to those employed above lead to the results given below; the geometrical relation $\cos \mathrm{OAD} = \cos\theta \cos\phi$ is required, the factor $\cos\phi$ converting V into its component V_x along Ox.

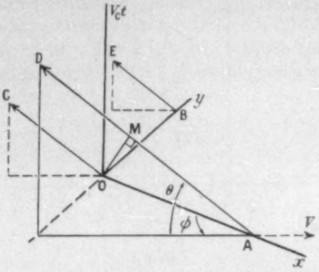


Figure 7.

$$V_x V_x' = V_c'^2, \qquad \ldots\ldots (19)$$

$$V_x = V\cos\phi = \xi_0/2\tau_{ex}, \qquad \ldots\ldots (20)$$

$$V_x' = V'/\cos\phi = \xi_0/\tau_{0x}, \qquad \ldots\ldots (21)$$

$$\tau'^2 - \tau^2 = (\xi_0^2 - 2V_x\xi_0\tau)/V_c'^2. \qquad \ldots\ldots (22)$$

In addition (14) and (15) still apply, V and V' being quantities that would be obtained directly by applying the formulae for a one-dimensional ground to records of two receivers in line with the drift V. τ_{ex} and τ_{0x} are written for the values of τ_e and τ_0 deduced from the records of the receivers on the x axis. A similar set of equations holds for the y axis, with the appropriate value η_0 for the receiver spacing in place of ξ_0, and with $\sin\phi$ replacing $\cos\phi$. Also we have

$$\tan\phi = V_y/V_x = V_x'/V_y' \qquad \ldots\ldots (23)$$

which, if $\eta_0 = \xi_0$, leads to the simple result

$$\tan\phi = \tau_{ex}/\tau_{ey} = \tau_{0y}/\tau_{0x}.$$

It is interesting to note that the equations for direction of V are not complicated by random changes; they are just the equations that would be expected in the special case considered in § 3.

We conclude from the results of this section that fading records taken at three receivers placed at corners of a right-angled triangle will be sufficient to enable us to solve our problem completely.

§ 6. GEOMETRY OF THE ρ-SURFACE

(i) *One-Dimensional Ground*; *Relations arising from Optimum Cross-correlation*

In this section we shall show how the results of equations (14) and (15) may be arrived at differently by a consideration of the geometry of the space–time correlation surface $\rho(\xi, \tau)$. As previously mentioned contours on the ρ-surface will be taken as ellipses of the form shown in Figure 4 (*c*).

Suppose two fading records are made at two points separated by a distance ξ_0, and it is required to find the time lag τ_0 which would produce the best cross-correlation between them. The problem can be solved by drawing a line on Figure 4 (*c*) through the point $(\xi_0, 0)$ parallel to the τ-axis, and determining the value of τ_0 where it touches one of the elliptical contours ρ_0. This value ρ_0 will then be the maximum value of the cross-correlation; let t_0 be the value of τ where this ellipse cuts the τ axis, so that t_0 is the time lag for *auto-correlation* required to give a value of correlation equal to this maximum *cross-correlation* corresponding to a lag τ_0.

x_0 is the intercept on the ξ axis, and (ξ_1, τ_1) is the point at which the tangent parallel to the ξ axis meets the ellipse. This notation accords with the symbols used in the definitions of § 4 so that equations (6), (8), (9) and (11) may be restated as our mathematical definitions:

$$V_c' = x_0/t_0, \quad V = \xi_1/\tau_1, \quad V_c = x_0\tau_1, \quad V' = \xi_0/\tau_0. \qquad \ldots\ldots (24)$$

Let the ellipse be $\qquad A\xi^2 + B\tau^2 + 2H\xi\tau = 1.$

If it has a vertical tangent touching at (ξ_0, τ_0) and passes through $(0, t_0)$, then

$$A = \frac{t_0^2 + \tau_0^2}{t_0^2\xi_0^2}, \quad B = \frac{1}{t_0^2}, \quad H = \frac{-\tau_0}{t_0^2\xi_0}. \qquad \ldots\ldots (25)$$

And from the fact that it meets a horizontal tangent at (ξ_1, τ_1) and also passes through $(x_0, 0)$, we may deduce

$$\tau_1^2 = \frac{A}{AB - H^2}, \quad \xi_1 = -\frac{H\tau_1}{A}, \quad x_0^2 = \frac{1}{A}. \qquad \ldots\ldots (26)$$

Equations (24), (25) and (26) enable us to express the various velocities in terms of A, B and H or in terms of our observables ξ_0, τ_0 and t_0:

$$\left.\begin{aligned} V_c' &= \left(\frac{B}{A}\right)^{\frac{1}{2}} = \frac{\xi_0}{(t_0^2 + \tau_0^2)^{\frac{1}{2}}}, \quad V = -\frac{H}{A} = \frac{\xi_0\tau_0}{t_0^2 + \tau_0^2}, \\ V_c &= \frac{[AB - H^2]^{\frac{1}{2}}}{A} = \frac{\xi_0\tau_0}{t_0^2 + \tau_0^2}, \quad V' = -\frac{B}{H} = \frac{\xi_0}{\tau_0}. \end{aligned}\right\} \qquad \ldots\ldots (27)$$

From (27) it is readily shown that

$$VV' = V_c'^2 = V_c^2 + V^2 \qquad \ldots\ldots (28)$$

which is equivalent to (14) and (15). In addition we have an alternative method of finding V_c', and hence V and V_c, because we have taken t_0 as observable;

comparison of the expression for V_c' in (27) with (16) shows that the connection with the previous method is given by

$$\tau_s^2 = t_0^2 + \tau_0^2. \qquad \ldots\ldots (29)$$

This relation may be demonstrated more directly by showing that the times τ_s, t_0 and τ_0 are proportional to the sides OA, OM and AM of the right-angled triangle OMA in the $(x, V_c t)$ diagram of Figure 4 (*a*).

(ii) *Other Methods*

The time τ_e is defined so that $\rho(0, \tau_e) = \rho(\xi_0, \tau_e)$. The ellipse $\rho = \rho_e$ that passes through points $(0, \tau_e)$ and (ξ_0, τ_e) is shown in Figure 5 (*c*). The mid-point of the horizontal chord between these points must lie on the same radial as the point of contact of the horizontal tangent. But we have seen that such a radial has slope $1/V$, so that

$$V = \xi_0 / 2\tau_e. \qquad \ldots\ldots (17)$$

We can also prove (18) from the geometry of an ellipse; if $A'\xi^2 + B'\tau^2 + 2H'\xi\tau = 1$ is the ellipse passing through (ξ_0, τ) and $(0, \tau')$, then $B'\tau'^2 = A'\xi_0^2 + B'\tau^2 + 2H'\xi_0\tau$ or $\tau'^2 - \tau^2 = (A'\xi_0^2 + 2H'\xi_0\tau)/B'$. Now since all correlation ellipses have been assumed similar, we may use ratios A/B and $2H/B$ for the ellipse $\rho = \rho_0$ discussed in (i) above. From (27) $A/B = 1/V_c'^2$ and $2H/B = -2/V' = -2V/V_c'^2$, and using these values in the expression above we get (18).

(iii) *Extension for Two-Dimensional Ground*

As R must now be regarded as a function of three variables, x, y, t, we have to consider the corresponding correlation function

$$\rho(\xi, \eta, \tau) = \frac{\overline{\{R(x+\xi, y+\eta, t+\tau) - \overline{R}\}\{R(x, y, t) - \overline{R}\}}}{\overline{\{R(x, y, t) - \overline{R}\}^2}}. \qquad \ldots\ldots (30)$$

The equivalent of our previous assumption of ellipses for lines of constant $\rho(\xi, \tau)$ is the assumption that surfaces of constant $\rho(\xi, \eta, \tau)$ are concentric ellipsoids centred on the origin of the coordinates. We also assume that the ground pattern is statistically isotropic, which implies that each correlation ellipsoid intersects planes perpendicular to the τ axis in circles.

Now we can consider in the same way as before the elliptical sections in the vertical planes $\eta = 0$, $\xi = 0$, and the vertical plane containing the direction of drift. Analysis of the latter section leads to (28), and to expressions for the velocities in terms of observations at points in line with the drift. Since this would not in general be of immediate use we must consider the other sections.

Figure 8 shows the vertical sections of the ellipsoid $\rho = \rho_{0x}$ corresponding to the maximum correlation between records of the two receivers on the x-axis. The $\eta = 0$ section therefore meets a vertical tangent at $(\xi_0, 0, \tau_{0x})$ and we have $V_x' = \xi_0 / \tau_{0x}$ as in (12); also, following the previous treatment, x_0 and t_{0x} are the intercepts on the ξ and τ axes, and $(\xi_1, 0, \tau_{1x})$ is the point of contact of the horizontal tangent.

The point (ξ_2, η_2, τ_2) where a horizontal plane is tangential to the ellipsoid gives

$$\tan\phi = \eta_2/\xi_2, \quad V = (\xi_2^2 + \eta_2^2)^{\frac{1}{2}}/\tau_2,$$

i.e. the components are $V_x = \xi_2/\tau_2$, $V_y = \eta_2/\tau_2$ as may be seen from the definition of V. Because the ellipsoid has circular horizontal sections it may be shown that the component V_x is also given by $V_x = \xi_1/\tau_{1x}$ which means that in this case

V_x can equally well be defined as the movement *along the x axis* that gives a minimum speed of fading.

Similar considerations apply to the receivers on the y axis which show a maximum cross-correlation $\rho = \rho_{0y}$ when the time lag is τ_{0y} and we obtain an elliptical section in the $(\eta\tau)$ plane with vertical and horizontal tangents at $(0, \eta_0, \tau_{0y})$ and $(0, \eta_1, \tau_{1y})$ and with intercepts y_0, t_{0y}. However, since all sections in this plane will be similar figures, we have scaled down the section in Figure 8 by a factor $k = x_0/y_0$ so that the diagram may represent for simplicity the sections of one ellipsoid having a circular cross section of radius x_0 in the $(\xi\eta)$ plane.

The remaining relations $V_c' = x_0/t_{0x}$, $V_c = x_0/\tau_2$ are straightforward statements of the definitions.

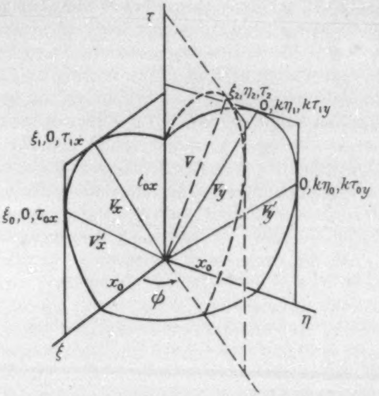


Figure 8. Sections of the ellipsoid $\rho = \rho_{0x}$

It is only necessary to consider the coordinate geometry of the ellipsoid in a way similar to that used for the ellipse in the case of one-dimensional ground to demonstrate from the mathematical definitions the results (19) to (23) of § 5. Also in (27) the expressions for V and V' in terms of t_0 and τ_0 are valid with suffixes x after these four symbols; the expression for V_c' is similarly valid, but the equation for V_c must be changed so that it is equivalent to

$$V_c^2 = V_c'^2 - V_x^2 - V_y^2, \qquad \text{......(31)}$$

a relation which follows from (14) and the fact that V_x, V_y are components of V.

§ 7. DISCUSSION OF RESULTS AND ANALYSIS OF RECORDS

We now turn to the way in which the results of the previous sections can be applied to the analysis of a set of three fading records. We are usually most interested in a determination of the velocity of drift V of the diffraction pattern since this provides evidence of steady motion somewhere in the ionosphere. But we may also be interested in the values of V_c as a measure of random changes, and in the factor which defines the fineness of structure of the diffraction pattern because these quantities have obvious importance in the study of fading.

The method of calculating V will depend on whether facilities or time permit a computation of auto- and cross-correlations of the records for a sufficient number of values of τ. Sometimes the time lags, τ_{0x} and τ_{0y}, for the best cross-correlation between records of receivers on the x and y axes respectively can be estimated fairly accurately by eye, without having to find ρ. If this is the case it is simplest to find the direction of V straight away from the equation $\tan\phi = \xi_0\tau_{0y}/\eta_0\tau_{0x}$ and to calculate V' from $1/V'^2 = \tau_{0x}^2/\xi_0^2 + \tau_{0y}^2/\eta_0^2$. These equations follow from (23), (21) and the corresponding expression for V_y'. We then need to obtain V from V'. If we recall the relation between the velocities $VV' = V_c'^2$ (equation (15)) we see that it is sufficient to find V_c'. As this quantity is the ratio of the speed of fading S to the factor G of the fineness of the ground pattern, it may be found from the auto-correlogram and the value of cross-correlation with no time shift (see equation (16)). Should we find V_c' approximately equal to V' we notice that either of these two quantities gives the approximate magnitude of V; this is quite a likely occurrence if, as has been supposed, the time lags of the records can be estimated visually.

If we perform a correlation analysis of the records we have the possibility of finding the time shift for optimum cross-correlation and, proceeding as in the previous paragraph, first finding V' and V_c' and then deducing the value of V. However, there are two alternative methods which give the components of V directly. The first method requires several values of both the auto- and cross-correlation so that the shift τ_e for which the correlations are equal may be found (see equation (20)). As a second alternative, a rather more thorough analysis will enable the graphical method described in § 5 (iv) to be used, and then the best values of V_c' as well as of the components of V will be obtained (see equation (22)). The value of the factor G is obtainable from V_c' and the speed of fading S by equation (7), $V_c' = S/G$.

In order to obtain the characteristic velocity V_c (which is a measure of the random changes) we must first find V_c' and the magnitude and direction of V, or V_x and V_y, and use equation (14) or (31). If V is large compared with V_c, so that V_c' becomes practically equal to V we may say that the drift is the predominant cause of fading. On the other hand if V_c is larger than V the fading is mainly due to random changes, and V becomes more difficult to find.

It is important in these calculations to perform some checks on the basis of the theory by comparing the two values for V_c' obtainable from the two pairs of stations. For if our assumption of an isotropic ground pattern holds the values should be equal. We should also obtain a third value of V_c' by exactly similar analysis of the third pair of records given by receivers A and B, and only if all three values agree can we be sure of isotropy. If the values are not in agreement a modified method of calculation is required which will be discussed in § 8.

One further point in these calculations is that is it not essential to evaluate the correlation function fully in order to find a maximum or equality of correlation; in practice it may be found that other methods of measuring correlation are simpler.

§ 8. LIMITATIONS AND EXTENSION OF THE THEORY

In order to discuss the nature of the assumptions involved in the analysis, we return to a further consideration of the correlation function $\rho(\xi, \tau)$ for a one-dimensional ground (Figure 4 (*c*)). Contours of constant ρ have been assumed to be ellipses. There is no obvious physical reason why this should be so, and it is conceivable that the contours might be of any shape whatever. However, the ρ-surface certainly has a maximum at the origin, where the correlation is necessarily unity. Now it is a theorem in analysis that sufficiently near a maximum or minimum, any surface has elliptical contours provided the principal radii of curvature are not infinite or zero. Consequently our assumptions will normally be valid in the limit for sufficiently small values of ξ and τ.

In considering the problem for a two-dimensional ground, we have assumed up to the present that the diffraction pattern was statistically isotropic so that at a given instant the auto-correlation function $\rho(\xi)$ as a function of distance is the same in all directions over the ground. This assumption was only made in order to simplify the analysis, and there is no difficulty in principle in extending the results to cover the possibility of a non-isotropic ground pattern. From the preceding paragraph it will be seen that the analysis will be made one stage more general if we allow the contours of equal correlation on the ground to have the form of ellipses. This defines two principal directions on the ground which we may conveniently take as axes of x and y. These axes will not, of course, have any relationship to the lines along which the observing receivers are placed. We now have two principal values of G, measured along the two axes, which we denote by G_x and G_y, and two values of V_c', defined by $V_{cx}' = S/G_x$ and $V_{cy}' = S/G_y$. The quantities S, S_c and V remain uniquely defined and require no generalization. An extension of the previous analysis enables all the above quantities to be found from fading records taken with three receivers placed at the corners of a right-angled triangle of arbitrary orientation. Since this extension involves no new principles it will not be given here.

A non-isotropic ground pattern of the type discussed above will be expected to occur whenever some direction in the ionosphere is specially distinguished. For example, in reflection at oblique incidence, the plane of propagation defines a special direction, and even at vertical incidence a special direction may be provided by the direction of the drift, or by the direction of the Earth's magnetic field.

Finally, we should point out that our analysis is only applicable on occasions when statistically stationary records of sufficient length are obtainable.

ACKNOWLEDGMENTS

The work here described forms part of a programme of Radio Research at the Cavendish Laboratory, supported by a grant from the Department of Scientific and Industrial Research.

We wish to thank our colleague Mr. S. A. Bowhill for helpful suggestions, particularly in connection with the methods of analysing records. We are also indebted to the Department of Scientific and Industrial Research for the individual maintenance allowances awarded to each of us.

REFERENCES

Mitra, S. N., 1949, *Proc. Instn. Elect. Engrs.*, Pt. III, **96,** 441.

Munro, G. H., 1948, *Nature, Lond.*, **162,** 886.

Pawsey, J. L., 1935, *Proc. Camb. Phil. Soc.*, **31,** 125.

Ratcliffe, J. A., 1948, *Nature, Lond.*, **162,** 9.

Ratcliffe, J. A., and Pawsey, J. L., 1933, *Proc. Camb. Phil. Soc.*, **29,** 301.

Les renforcements brusques des ondes très longues

PAR R. BUREAU
Laboratoire National de Radioélectricité, Bagneux, France

Communicated by J. A. Ratcliffe; MS. received 26th July 1949. *Paper read at Physical Society Conference at Cambridge, 14th–16th July* 1949

ABSTRACT. During the last twenty years, enhancements of the mean level of atmospherics, on 11 km. wavelength, have been associated with some other phenomena such as simultaneous fade-outs of the field intensity of stations on decametric wavelengths, perturbations of terrestrial magnetism or Wolf numbers relative to solar activity.

More recently a search has been made for occasions on which the association between the different phenomena is not so close. It has been found that the usual fade-out on short waves does not occur in the same manner on all stations recorded; sometimes, indeed, an enhancement of short duration takes place at the beginning of the expected fade-out. This enhancement is probably due to the reception of a radioelectric radiation from the sun.

The author gives some examples of such phenomena which can be used for a better forecasting of radioelectric disturbance due to abnormal solar activity.

LES renforcements brusques des ondes très longues (kilométriques et myriamétriques) et tout particulièrement ceux du niveau des atmosphériques ont joué et jouent encore un rôle essentiel dans l'étude des perturbations ionosphériques à début brusque (P.I.D.B.).

Cette étude a débuté par la recherche des synchronismes. Ainsi a-t-on groupé autour des P.I.D.B. des manifestations très diverses; ce travail fut facilité par la brusquerie avec laquelle s'installent les phénomènes. Donnons-en deux exemples parmi beaucoup d'autres:

le 21 *février* 1942 (figure 1*) début brusque à 13 h. 26 sur le niveau moyen des atmosphériques et, simultanément, crochet sur la déclinaison du magnétisme terrestre (exemple fameux, synchrone avec une éruption chromosphérique intense se détachant sur le bord E du soleil);

le 28 *mai* 1936 à 14h. et à 18h. T.U., (figure 2) disparition des échos ionosphériques à Huancayo et en synchronisme deux renforcements des atmosphériques à Rabat.

Autre mode de comparaison, mais à une autre échelle: établir la liste ininterrompue des P.I.D.B. Ce travail s'accommode très bien de l'enregistrement des atmosphériques sur 11,000 mètres. On met ainsi en évidence l'influence du cycle solaire undécennal; la figure 3 donne le nombre annuel des renforcements pendant 20 ans; on y retrouve deux maximums successifs correspondant aux maximums d'activité solaire.

On entre maintenant dans une nouvelle période de la recherche; celle qui se base sur l'étude des divergences. Ici encore deux méthodes s'offrent à nous:

(*a*) une méthode statistique où l'on relève le pourcentage d'apparitions synchrones des diverses manifestations rattachables aux P.I.D.B.

(*b*) l'étude d'exemples individuels.

* For all photographs see group of Plates following p. 128.

Ce sont des exemples de cette dernière méthode que nous donnerons ci-après en choisissant quatre cas qui se rattachent aux rayonnements radioélectriques solaires. Il était indiqué de rechercher si certains de ces phénomènes nouveaux, dont l'apparence rappelait si étrangement les P.I.D.B., n'étaient pas en liaison avec elles, tout au moins partiellement. C'est ce qu'ont fait Appleton et Hey qui ont montré qu'il y avait une concordance remarquable entre les deux phénomènes pendant une période d'activité solaire intense qui s'est déroulée du 30 janvier au 14 février 1946. Or, comme il a été signalé par Appleton (1945, 1946), il peut y avoir un antagonisme dans les actions sur l'ionosphère, provoquées au même moment, les unes par les rayonnements ultra-violets, les autres par le rayonnement électromagnétique sur ondes métriques et décamétriques. Je verse au dossier de cette enquête quatre documents dans lesquels semble bien se manifester cet antagonisme. Le premier, que j'ai déjà signalé (Bureau 1947), relatif au 8 février 1946 et trois autres relatifs au 28 décembre 1948, 5 avril 1949 et 7 mai 1949.

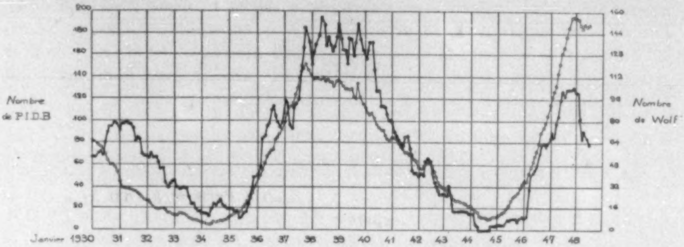


Figure 3.

—○—○—○—○—○—○— Moyennes mobiles annuelles de l'activité solaire (nombre de Wolf).

—●—●—●—●—●—●— Moyennes mobiles annuelles du nombre des P.I.D.B.

Remarque. Les moyennes concernent les 12 mois précédant chaque point.

8 *février* 1946 (figure 4)

La courbe supérieure est celle du niveau moyen des atmosphériques sur 11,000 mètres de longueur d'onde le 8 février 1946 de 09 h. à 13 h. T.U. La courbe inférieure donne le champ de Genève sur 48,66 mètres. On trouve là un exemple particulièrement net du renforcement sur 11,000 mètres accompagné d'un affaiblissement sur 48,66 mètres. Mais ce qui rend cet enregistrement particulièrement précieux c'est que l'évanouissement de Genève est précédé d'un renforcement brutal d'une durée au plus égale à quelques secondes (la plume n'a pas eu le temps d'inscrire un trait continu dans la montée). Il faut sans doute voir là l'exemple d'un sifflement que certains auteurs ont signalé parfois au début d'un évanouissement brusque. Le mécanisme tel que l'envisageait Appleton serait le suivant : au début de la perturbation le rayonnement radioélectrique solaire atteint la terre, après la traversée des régions E et F de l'ionosphère avant que le rayonnement ultra-violet n'ait tendu un écran ionisé dans la région D. Très peu de temps après, l'écran est tendu et l'évanouissement se produit. Le raisonnement d'Appleton s'applique à des fréquences comprises entre 10 et 30 Mc/s. (ondes métriques et décamétriques).

28 décembre 1948 (figure 5)

La courbe supérieure reproduit l'enregistrement des atmosphériques à Bagneux sur 11,000 mètres de longueur d'onde le 28 décembre 1948 de 09 h. à 16 h. T.U. La figure inférieure a trait à l'enregistrement à Bagneux du champ de l'émetteur de Leipzig sur 30,80 mètres. Le renforcement des atmosphériques n'est pas accompagné d'un évanouissement sur ondes décamétriques, mais il accuse une montée rapide, synchrone avec celle des atmosphériques sur 11,000 mètres. Cette montée cesse très peu de temps après et la plume redescend au niveau précédent. Aucun évanouissement à début brusque n'est signalé en Europe ni en Amérique le 28 décembre. Deux appareils totalement distincts ont enregistré à Bagneux la montée sur 11,000 mètres*. Se trouve-t-on là en présence du même phénomène que ci-dessus ? On ne peut l'affirmer d'une manière absolue mais il était bon de faire connaître cet exemple.

5 avril 1949 (figure 6)

La figure 6 a trait à une partie de la journée du 5 avril 1949, à cheval sur le jour et la nuit. Le soleil s'est couché à 18 h. 26 à Paris et à 18 h. 51 à Rabat. A la fin de la journée ont apparu, séparées d'à peu près une heure, deux P.I.D.B. particulièrement importantes sur 11,000 mètres. La première colonne de la figure reproduit trois enregistrements distincts d'atmosphériques sur environ 11,000 mètres à Bagneux, ainsi qu'un quatrième relevé à Rabat. La seconde colonne reproduit pour la même période trois enregistrements d'ondes décamétriques effectués à Bagneux et correspondant aux émissions de (*a*) Leipzig sur 9,732 kc/s. (30,80 mètres), (*b*) et (*c*) Washington WWV sur 2,000 kc/s. (15 mètres). (La première courbe de Washington correspond au champ de l'onde porteuse, la seconde reproduit l'onde porteuse plus la modulation.) En bas de la seconde colonne, on trouve un extrait du goniogramme à secteur étroit de Trappes (environ 25 km. à l'ouest de Bagneux).

La première P.I.D.B. est du type habituel; début à 15 h. 18, maximum à 15 h. 29, fin vers 16 h. 05. Elle est tout aussi nette à Rabat qu'à Bagneux, après une période de brouillages qui s'y termine, heureusement, un peu avant 15 h. Le goniogramme de Trappes y est très sensible, surtout pour la direction du foyer principal, c'est à dire vers le Sud. Le champ de Leipzig semble complètement indifférent à la P.I.D.B. Le champ de Washington accuse un évanouissement dont le début semble bien concorder avec le départ du renforcement sur 11,000 mètres. La descente se termine à 15 h. 27. Des évanouissements ont été constatés sur le trafic commercial en particulier à la station de Radio-France à Villecresnes où l'on a signalé à 15 h. 15 un important évanouissement sur le trafic de New-York. Le C.R.P.L. signale un certain nombre d'affaiblissements mais indique comme point de départ 15 h. 22.

Cette première P.I.D.B. respecte les caractères généralement admis relatifs à l'effet inverse des ondes décamétriques et des ondes kilométriques. Tout au plus peut-on signaler que l'évanouissement de Washington à Bagneux est précédé d'une pointe de très courte durée particulièrement visible sur l'enregistrement mixte (onde porteuse et modulation). Ce renforcement fait songer à celui du 8 février 1946 décrit plus haut. Notons aussi que l'évanouissement signalé par le C.R.P.L. est de 4 minutes plus tardif que les heures de départ précitées.

---

* Un observateur se trouvait présent devant les appareils au moment de ce phénomène et a pu constater le synchronisme relaté ci-dessus.

La deuxième P.I.D.B. est apparue brutalement à 16 h. 25. Elle se compose vraisemblablement de deux P.I.D.B. superposées, mais le partage entre l'une et l'autre est simplement indiqué par un léger palier des courbes des atmosphériques enregistrées à Bagneux. La séparation paraît plus nette à Rabat. Les heures de départ sont à 16 h. 25 et 16 h. 35. La descente pour les deux P.I.D.B. commence à 16 h. 54 pour se terminer vers 17 h. 30 dans le minimum qui précède le coucher du soleil, minimum très marqué sur le goniogramme. Le départ de la première partie de la deuxième P.I.D.B. est marqué sur le goniogramme, à la fois sur le foyer Sud et sur le foyer Ouest-Nord-Ouest.

Voici maintenant ce que l'on a constaté sur les évanouissements. Sur Leipzig s'observe une montée continue oscillant autour d'une courbe médiane qui donne l'apparence d'une montée logarithmique et qui correspond sans doute à la diminution de l'absorption. Insensible, comme nous l'avons dit plus haut à la P.I.D.B. de 15 h. 18, le champ passe par un maximum entre 16 h. 32 et 16 h. 35. Il accuse ensuite une chute rapide qui se termine, avant d'avoir atteint un évanouissement total, et cela légèrement après le renforcement brusque sur 11,000 mètres. Sur Washington la courbe de Bagneux accuse une descente en deux parties avec un palier correspondant à celui inscrit sur 11,000 mètres. La chute de Leipzig commence au moment où finit ce palier. Des évanouissements sont signalés à 16 h. 38 à New-York et 16 h. 22 à Washington.

Après le coucher du soleil le goniogramme d'atmosphériques signale une montée nocturne qui concerne tout d'abord une puissante source aux environs du Sud. Mais comme on le sait la nuit il n'y a pas de P.I.D.B.

Dans cet exemple nous trouvons des anomalies seulement sur le champ de Leipzig qui n'éprouve pas d'évanouissement dans la première P.I.D.B. et qui, de plus, accuse un renforcement provisoire entre la première et la seconde partie de la deuxième P.I.D.B. Le champ de Leipzig s'affaiblit seulement au moment où apparaît la seconde partie de la dernière P.I.D.B.

7 *mai* 1949 (figure 7)

La courbe supérieure représente l'enregistrement des atmosphériques sur 11,000 mètres à Bagneux de 09 h. à 16 h. T.U. le 7 mai 1949. La courbe centrale donne l'enregistrement à Bagneux du champ de Washington WWV sur 15 mètres, la courbe du bas l'enregistrement de Moscou sur 26,60 mètres. Deux P.I.D.B. se manifestent à 10 h. 45 et 13 h. 45 T.U. Sur la première on retrouve un très rapide et bref renforcement des ondes décamétriques de WWV précédant l'évanouissement habituel. Un observateur se trouvait présent devant les enregistreurs et a constaté la chronologie des diverses phases du phénomène, et en particulier la simultanéité de la montée de WWV avec la montée des atmosphériques sur 11,000 mètres. La seconde P.I.D.B. est du type normal, l'évanouissement des ondes courtes concordant avec le renforcement des ondes longues.

Ces quelques exemples ont seulement pour but de mettre en lumière un phénomène qui peut être très important â l'avenir surtout dans les prévisions d'orages magnétiques ou d'orages ionosphériques. On sait en effet que dans un certain nombre de cas on peut associer une P.I.D.B. (due à un rayonnement ondulatoire solaire) au déclenchement ultérieur d'un orage magnétique (dû à un rayonnement corpusculaire), ce qui entraine des troubles très profonds dans les radiocommunications, le délai qui sépare les deux phénomènes étant en général de 36 heures et pouvant varier d'environ 20 heures à environ 40 heures. On conçoit ainsi à quel point il est intéressant de rattacher ces phénomènes du rayonnement à une troisième sorte de rayonnement dû au soleil et qui est un rayonnement radioélectrique décelable sur les ondes métriques, décimétriques et même centimétriques. Il est donc assez curieux de signaler que les P.I.D.B. si spectaculaires du 5 avril ont été signalées aux services radioélectriques comme annonçant des perturbations probables de l'ionosphère dans un délai allant de 36 heures à 50 heures. Or, le 8 avril à 14 h. T.U. les liaisons par ondes décamétriques entre Paris et New-York ont été totalement impossibles.

On ne peut affirmer ici, qu'il y a une relation certaine de cause à effet entre telle P.I.D.B. et tel trouble ultérieur dans la propagation, mais il est certain que l'on peut dès maintenant envisager l'amélioration de messages d'alerte grâce à une meilleure connaissance des perturbations ionosphériques à début brusque et aux phénomènes qui s'y rattachent. Les enregistrements permanents sont en tout état de cause d'un secours inappréciable.

BIBLIOGRAPHIE

Appleton, E. V., 1945, *Nature, Lond.*, **156**, 534; 1946, *Phil. Mag.*, **37**, 73.
Bureau, R., 1947, *L'Onde Elect.*, **27**, 45.

Diffusion des Echos au Voisinage des Fréquences Critiques de F2

Par R. RIVAULT
Université de Poitiers

Communicated by J. A. Ratcliffe ; MS. received 26th July 1949. Paper read at Physical Society Conference at Cambridge, 14th–16th July 1949

ABSTRACT. From the study of ionospheric data recorded at Poitiers (46° N.) it appears that two sorts of scattered echoes take place in the neighbourhood of F2 critical frequencies.

The first one seems to be produced by the multiplication of magneto-ionic components. This scattering ends with ground sunrise; it exists during both quiet and perturbed ionospheric conditions; it is far more frequent and stronger when the nights are longer. Its existence seems to be connected with nocturnal thermal and convective disorder.

The second sort of scattering is mainly recorded during the second part of the night, after F2 penetration, at a virtual height of about 800 to 1,200 km. The sharp lower boundary of these reflections, which disappears a few moments after sunrise, leads to the postulation of an ionospheric G region, whose existence has been often discussed. This sort of scattering is connected with magnetically disturbed conditions and, chiefly, with meteoritic periods.

Dans le cas le plus général, on déduit de la théorie magnéto-ionique que, pour un certain niveau d'ionisation maximum de la région F2, trois fréquences critiques peuvent être enregistrées, correspondant aux relations :

$N\mathbf{e}^2/4\pi^2\mathbf{m}=f^2{}_{N\,\mathrm{max}}=f^2\mathrm{o}$ pour le rayon ordinaire,

$f^2{}_{N\,\mathrm{max}}=f\mathrm{x}1(f\mathrm{x}1-f_H)$ pour l'extraordinaire le plus intense,

$f^2{}_{N\,\mathrm{max}}=f\mathrm{x}2(f\mathrm{x}2+f_H)$ pour l'extraordinaire le moins intense, avec $f_H=\mathbf{e}H/2\pi\mathbf{m}$.

D'où : $f\mathrm{x}1-f\mathrm{o}=f_H f\mathrm{x}1/(f\mathrm{x}1+f\mathrm{o})\simeq f_H/2$ si $f\mathrm{o}$ et $f\mathrm{x}1$ sont grands devant f_H, et: $f\mathrm{x}1-f\mathrm{x}2=f_H$.

Les sondages à la verticale effectués à Poitiers (lat. 46° 34′ N.) comprennent habituellement les composantes ordinaire *fo* et extraordinaire *fx*1, espacées d'environ 0,6 Mc/s., soit $f_H/2$. La composante *fx*2 n'apparait que très rarement, généralement sur les réflexions d'ordre 2 et dans les premières heures de la matinée espacée de *fx*1 d'environ 1,2 Mc/s. (Figure 1*). L'on rencontre beaucoup plus fréquemment des composantes régulièrement ou irrégulièrement réparties dans l'intervalle de fréquence *fx*1–*fx*2. Chacune des composantes *fo* et *fx*1 peut d'abord se dédoubler (Figure 2), puis se multiplier en même temps que les échos perdent leur netteté (Figures 3, 5 et 6), pour, enfin, s'agglomérer en une masse diffuse (Figure 4). Quand ces composantes supplémentaires sont assez fines pour permettre des mesures, on remarque très souvent qu'il existe entre elles des écarts de fréquence privilégiés, de l'ordre de 0,3 (Figure 3) et de 0,15 Mc/s. (Figure 5), valeurs sous multiples de f_H.

A Poitiers, la composante verticale du champ magnétique terrestre est environ le double de la composante horizontale (0,42 et 0,20 gauss au sol), et l'on peut penser que, sur leur trajet, les ondes sont soumises temporairement à des conditions de propagation alternativement transversales ou longitudinales; ceci peut être dû à une variation soit du nombre de chocs ν, provoquée par une modification locale de la densité du milieu ou du nombre d'électrons, soit du champ magnétique terrestre H. Breit a montré, en effet, qu'il suffit d'une variation de H de l'ordre de 20γ pour faire tourner le plan de polarisation de l'onde de 90°. Etant donnée la précision bien inférieure des mesures possibles sur les enregistrements, il n'est pas surprenant que la valeur de f_H, déduite de celles-ci, n'en paraisse pas affectée: les écarts de fréquence des composantes homologues ordinaires et extraordinaires de la figure 5 sont, par exemple, de l'ordre de 0,6 Mc/s. Le rôle de la polarisation est mis en évidence par le fait que ces composantes multiples n'apparaissent pas simultanément sur les échos d'ordre un et ceux d'ordre supérieur; une analyse plus fine nécessiterait la mise en oeuvre de dispositifs expérimentaux de polarisation connue et variable.

L'examen de nombreux enregistrements analogues aux figures précédentes conduit à imaginer la diffusion des échos qui se manifeste à certaines époques, au moment de la pénétration de F2, comme le résultat de la multiplication du nombre des composantes magnéto-ioniques; c'est la même cause physique qui se traduit par ces différents aspects.

De l'étude des cas de diffusion enregistrés de juillet 1948 à avril 1949, entre 3,1 et 11,8 Mc/s., on peut tirer les conclusions suivantes:

(i) la diffusion des échos, au voisinage des fréquences critiques de F2, ne se produit nettement qu'entre le coucher et le lever du soleil;

(ii) la fréquence du phénomène, faible pendant l'été, croît à partir de septembre pour devenir quasi quotidienne en décembre, puis décroît ensuite pour n'être plus qu'accidentelle en mars;

(iii) elle se manifeste généralement dans la deuxième partie de la nuit mais peut débuter dans la première et d'autant plus tôt que les nuits sont plus longues. Elle prend fin au moment du lever du soleil au sol, un peu après le début de l'accroissement matinal des fréquences critiques;

(iv) les enregistrements à caractère diffus sont obtenus aussi bien dans des conditions magnétiques calmes que perturbées (figure 6); le phénomène ne semble pas modifié, non plus, au cours des périodes de grande activité chromosphérique solaire.

---

* For all photographs see group opposite p. 128.

La diffusion parait ainsi directement liée à la longueur de la nuit, comme si, entre autres causes possibles, cette durée favorisait le désordre thermique et électronique de l'ionosphère, détruisant la stratification diurne jusqu'au moment du lever solaire.

Il existe une autre sorte de reflexions diffuses, enregistrées à des hauteurs virtuelles de 800 à 1200 kilomètres, sur des fréquences immédiatement supérieures aux fréquences critiques de F2. Elles peuvent se produire au cours de la journée et sont alors faibles (figure 7). Mais elles sont parfois reçues d'une façon intense dans la deuxième partie de la nuit, quand les fréquences critiques de F2 sont basses; leur aspect particulièrement net conduit alors à imaginer l'existence d'une véritable région ionisée désignée par la lettre G (figure 8), existence niée par Bailey, Payne Scott et McCready à la suite d'expériences d'un caractère tout différent.

Ces réflexions nettes se produisent généralement sur des fréquences allant de 3 Mc/s. environ à 5,2 Mc/s. ; elles disparaissent, exceptionnellement, sur une fréquence plus élevée (figure 9). Un sondeur plus puissant serait nécessaire pour mettre en évidence les fréquences critiques de cette région. Elles prennent habituellement fin aux environs du lever solaire, comme la diffusion étudiée dans la première partie de cette note. Il arrive, du reste, que ces deux types d'échos diffus se confondent, la région G perdant progressivement son caractère propre pendant que les composantes magnéto-ioniques habituelles s'estompent dans une masse diffuse (figure 10).

Ces deux espèces de diffusion ne sont cependant pas toujours concomitantes, la présence de la région G coïncidant souvent avec des époques magnétiquement perturbées et, surtout, avec le passage des principaux essaims de météorites.

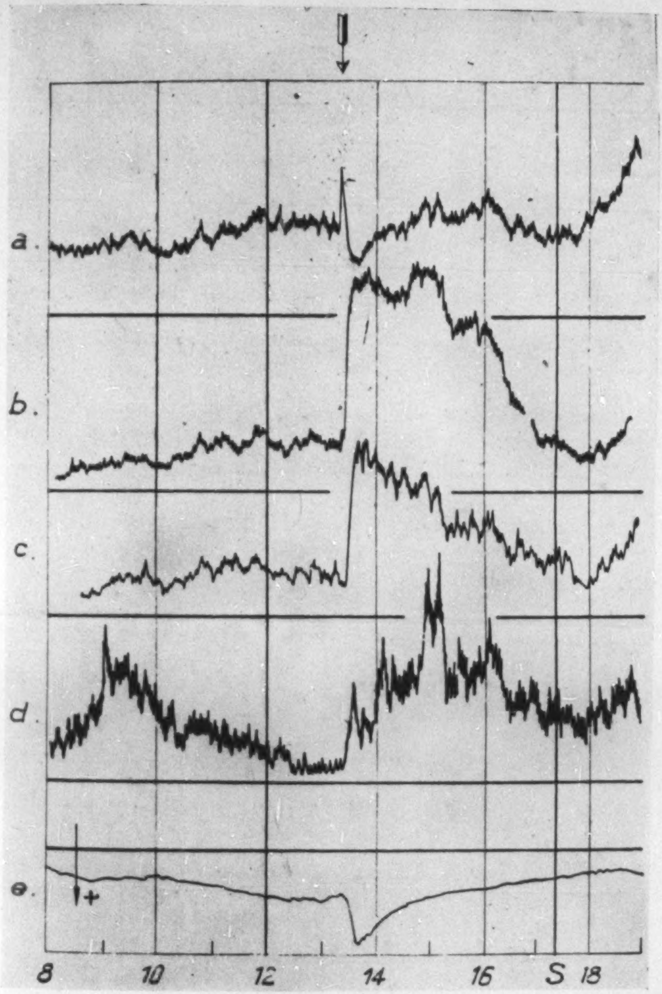


Figure 1. P.I.D.B. du 21 Février 1942.

Successivement à partir du haut:

a: Enregistrement des atmosphériques sur 23,000 m.
b, *c*: Enregistrement des atmosphériques sur 11,000 m. (deux appareils distincts).
d: Enregistrement des atmosphériques sur 5,000 m.
e: Enregistrement de la déclinaison magnetique à Chambon la Forêt.

To face page 128

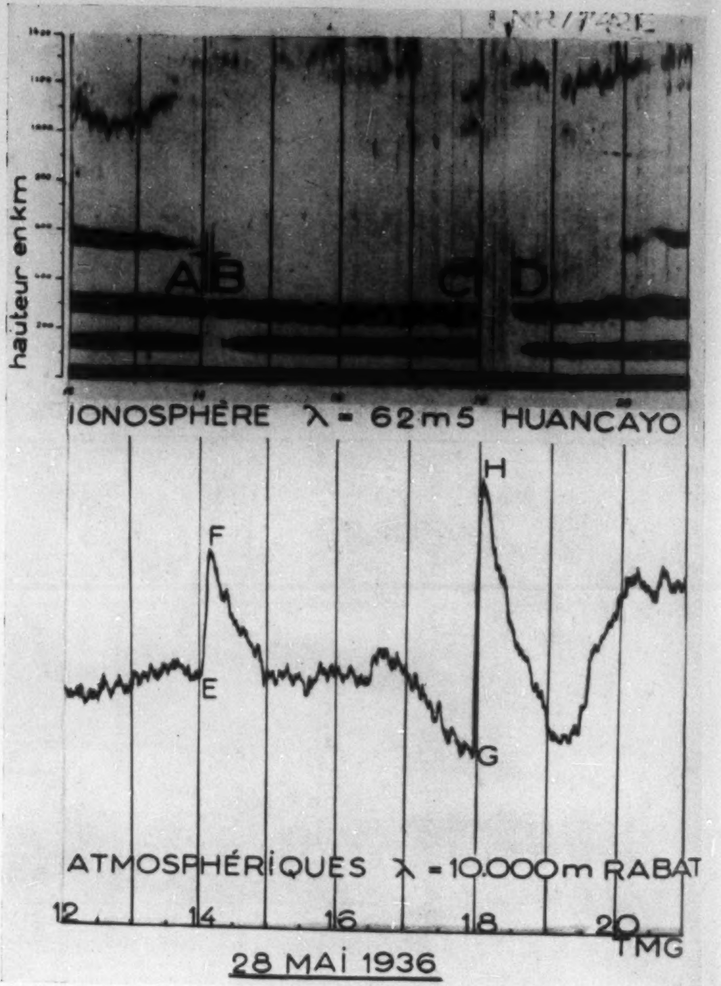

NR/742E
hauteur en km
A
B
C
D
IONOSPHÈRE λ = 62 m 5 HUANCAYO
H
F
E
G
ATMOSPHÉRIQUES λ = 10.000 m RABAT
12
14
16
18
20
TMG
28 MAI 1936

Figure 2.

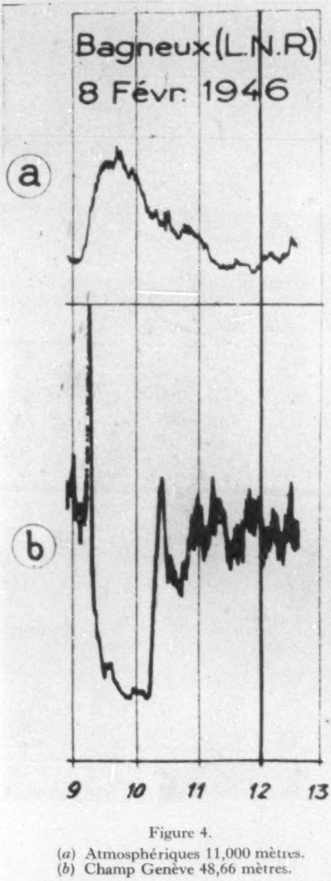


Figure 4.

(*a*) Atmosphériques 11,000 mètres.
(*b*) Champ Genève 48,66 mètres.

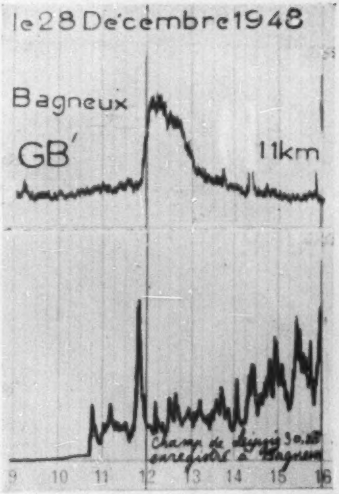


Figure 5.

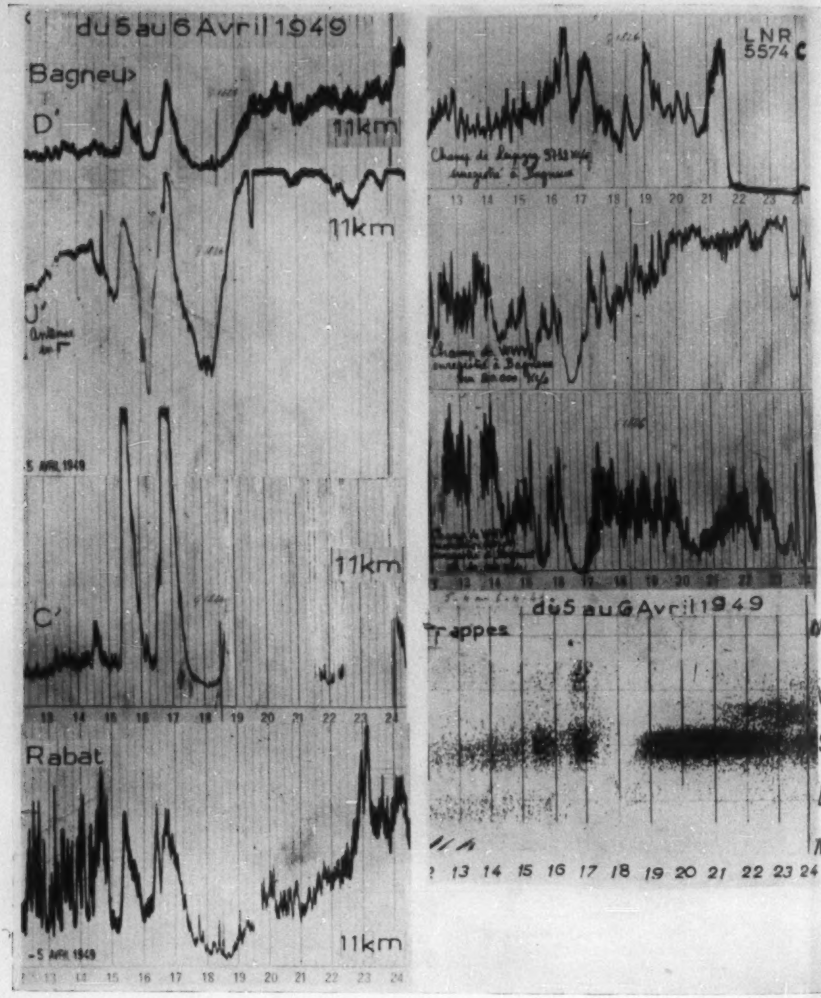

du 5 au 6 Avril 1949
Bagneux
D'
11km
U'
Antenne en F
5 AVRIL 1949
11km
C'
11km
13 14 15 16 17 18 19 20 21 22 23 24
Rabat
5 AVRIL 1949
11km
13 14 15 16 17 18 19 20 21 22 23 24
LNR 5574 C
enregistré à Bagneux
13 14 15 16 17 18 19 20 21 22 23 24
Champ de WWV enregistré à Bagneux
du 5 au 6 Avril 1949
Trappes
N
W
S
E
N
13 14 15 16 17 18 19 20 21 22 23 24

Figure 6.

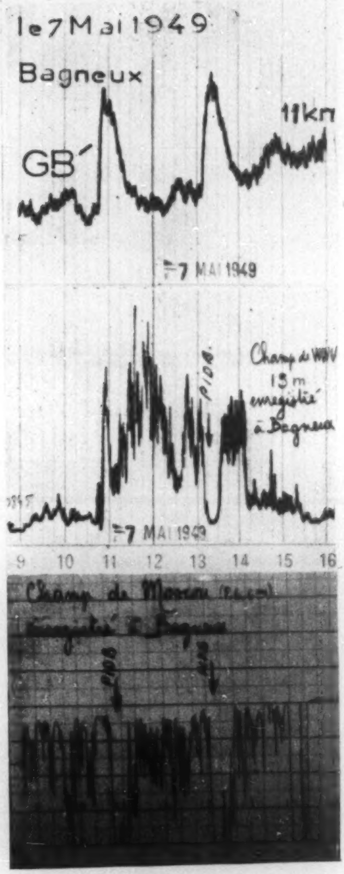

le 7 Mai 1949
Bagneux
GB'
11km
7 MAI 1949
PIDB
Champ de WWV
15 m
enregistré
à Bagneux
7 MAI 1949
9
10
11
12
13
14
15
16
Champ de Moscou
enregistré à Bagneux
PIDB

Figure 7.

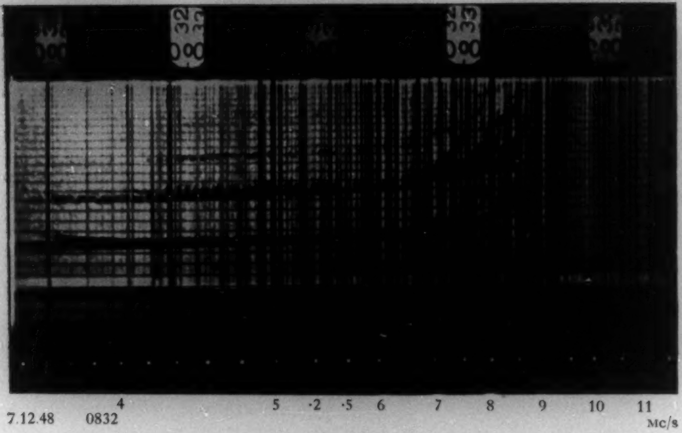


Figure 1.

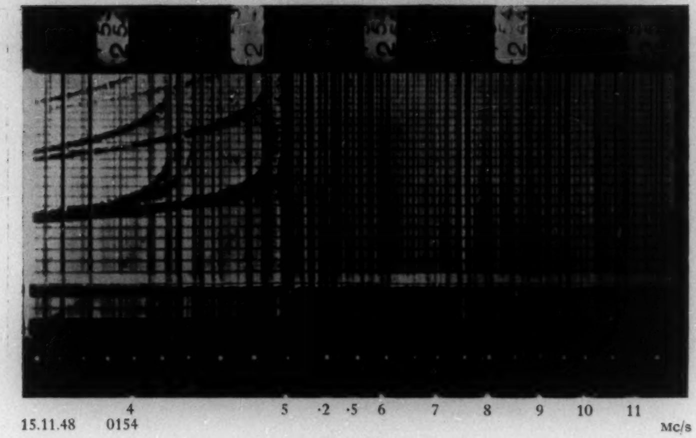
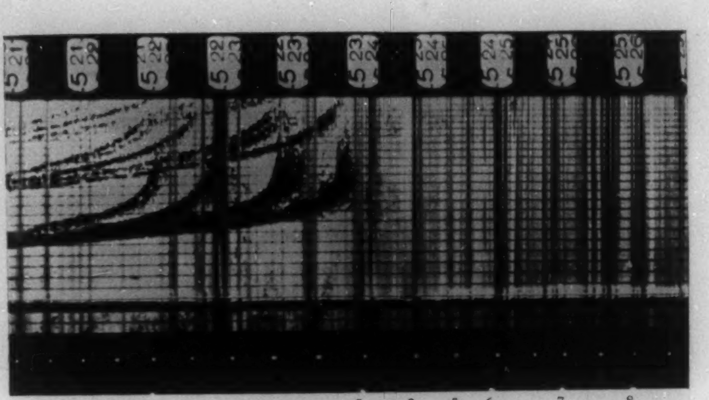
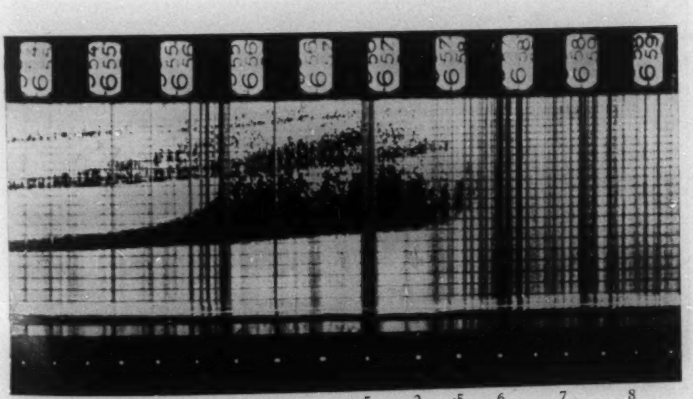


Figure 2.



17.10.48 0521

Figure 3.



16·10·48 0555

Figure 4.

PROC. PHYS. SOC. VOL. 63, PT. 2—B (R. RIVAULT)

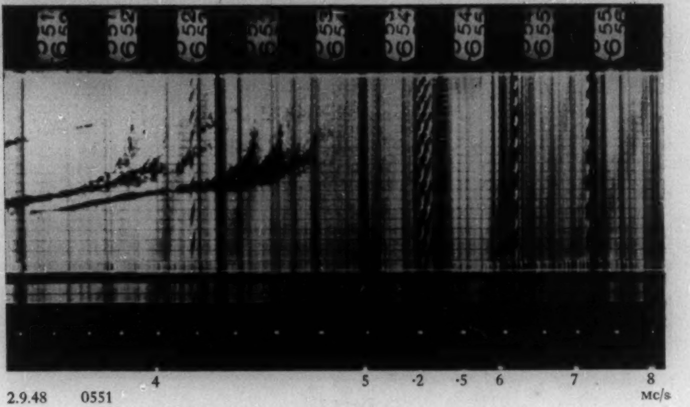


Figure 5.

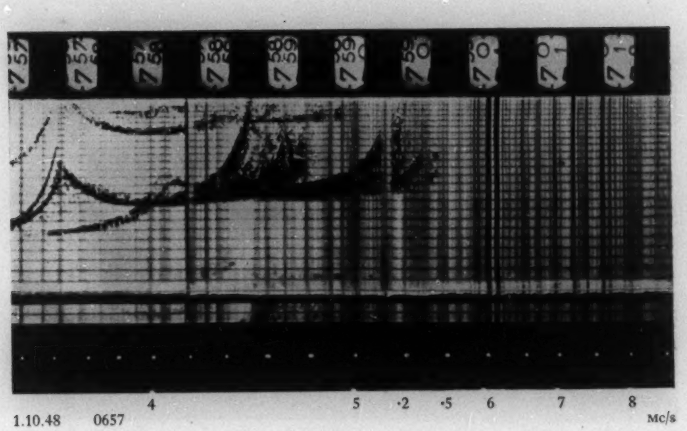


Figure 6.

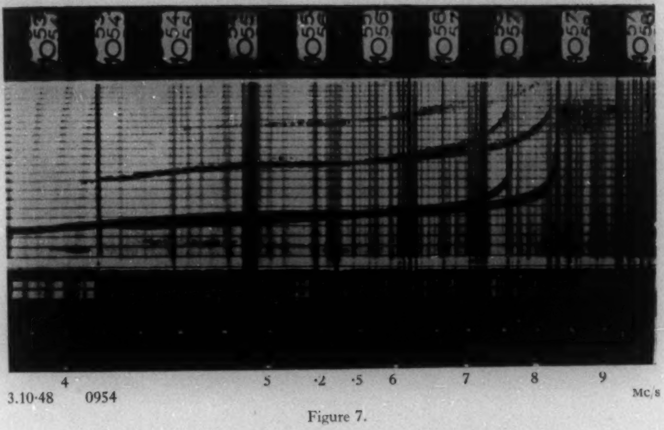


Figure 7.

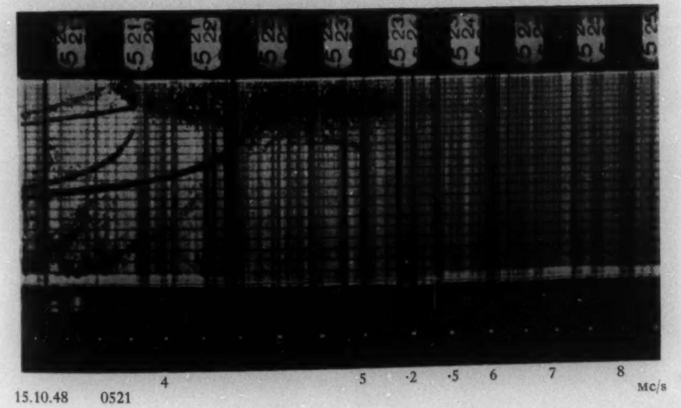


Figure 8.

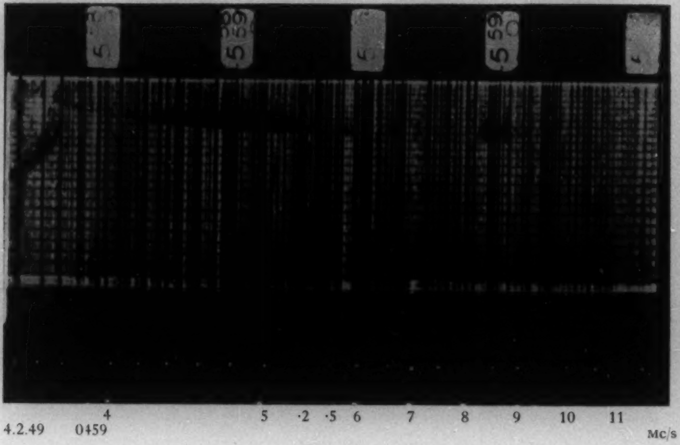


Figure 9.

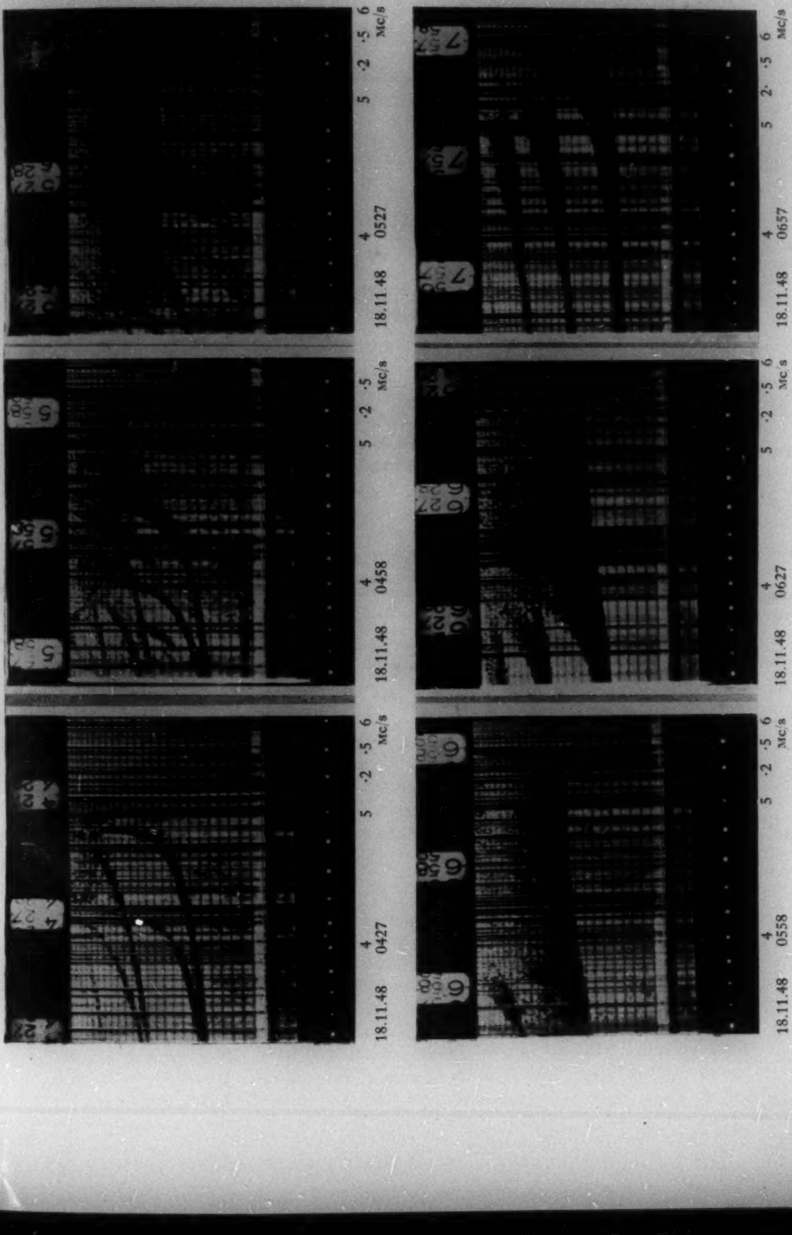

18.11.48
0527
4
5
·2
·5
6
Mc/s
18.11.48
0458
4
5
·2
·5
Mc/s
18.11.48
0427
4
5
·2
·5
6
Mc/s
18.11.48
0657
4
5
2·
·5
6
Mc/s
18.11.48
0627
4
5
·2
·5
6
Mc/s
18.11.48
0558
4
5
·2
·5
6
Mc/s

Figure 10.

Theoretical Considerations Regarding the Formation of the Ionized Layers

By D. R. BATES and M. J. SEATON
University College, London

MS. received 1st September 1949; read at the Summer Meeting of the Physical Society at Cambridge, July 1949

ABSTRACT. The detailed mechanisms involved in the formation of the E, F_1, F_2 and D layers by solar ultra-violet radiation are discussed. Use is made of the results of some recent calculations on the continuous absorption cross section of atomic oxygen and nitrogen, and of the evidence on the ionic composition of the layers that is provided by the analysis of the emission spectrum of the upper atmosphere during twilight. The uncertainties existing at present are emphasized.

§ 1. INTRODUCTION

For a proper theory of the formation of the ionized layers it is necessary to have detailed information on the incident flux of solar radiation, on the altitude distribution and chemical composition of the terrestrial atmosphere, and on the absorption cross-section curve associated with each constituent. Unfortunately the information available on these topics is far from complete and in consequence the development of the theory must proceed slowly and tentatively. In the present paper we confine ourselves to a discussion of the origin of the so-called ultra-violet layers (that is, the E, F_1, F_2 and D layers), and ignore the effects of corpuscles and meteors; for convenience, we treat the three upper, and the lowest, separately. Emphasis is laid on the existing uncertainties.

§ 2. THE E, F_1 AND F_2 LAYERS

2.1. Some years ago Bradbury (1938) suggested that the F_1 and F_2 layers might have a common origin. Bates (1949 a) has recently examined this hypothesis quantitatively and has found that it is consistent with the observational data. Indeed the conclusion that it *must* be largely correct is difficult to avoid as the ionization rate near the 300 km. level due to radiation capable of producing the F_1 layer can be shown to be sufficient *by itself* to account for the F_2 layer also. That there are two layers is simply due to the decrease of the effective recombination coefficient with altitude. The F_1 layer approximates to the standard Chapman type (1931 a) so that its maximum occurs near the level at which the concentration $n(X)$ of the atoms or molecules X responsible for the ionization is given by

$$n(X) = \cos\chi / HA, \qquad \ldots\ldots(1)$$

χ being the solar zenith angle, H the local scale height and A the relevant absorption cross section In contrast the F_2 layer lies above the region where there is a marked diminution of the intensity of incident radiation and is what may be termed a low attenuation layer, the principal characteristic of which is the relatively weak direct control over the ionization rate exerted by the elevation of the Sun—actually for an extreme zero attenuation layer the ionization rate depends only on the gas density at the level concerned. As far as can be judged the observed diurnal and seasonal variations are in agreement with the theory; analysis of the irregular fluctuations associated with solar activity shows, too, the expected correlation between the behaviour of the two layers (cf. Allen 1948, Bates 1949 a).

The atmosphere in the region above about 100 km. is generally considered to be composed mainly of molecular (or perhaps atomic) nitrogen and atomic oxygen. It is reasonable therefore to presume that photo-ionization of one or more of these is responsible for the F_1 and F_2 layers. There is some spectroscopic evidence which strongly suggests that the active constituent is not molecular nitrogen. N_2^+ ions resonate strongly in sunlight, emitting bands of the first negative system :

$$\left.\begin{array}{l} N_2^+ + h\nu \rightarrow N_2^{+\prime} \\ N_2^{+\prime} \rightarrow N_2^+ + h\nu. \end{array}\right\} \qquad \ldots\ldots(2)$$

In consequence the emission spectrum of the twilit upper atmosphere provides an extremely sensitive means of detecting their presence. By analysing observational data published by Dufay and Dufay (1947) Bates (1949 a) has shown that during twilight $[N_2^+]$ is only 10 per cm³ or less at the level of the F layers, which is far smaller than the total ion concentration. The simplest interpretation* is that ionization of molecular nitrogen occurs very slowly and is only of minor importance.

The suggested elimination of molecular nitrogen by the spectroscopic evidence mentioned is particularly useful as otherwise the lack of quantitative information on its absorption cross-section curve would be a serious hindrance to progress. Fortunately both the remaining constituents are atomic, and thus their photo-ionization can be investigated theoretically without great difficulty. Some early calculations on the cross sections were performed by Bates, Buckingham, Massey and Unwin (1939) and by Bates (1939). Adopting the results given by these, and making reasonable estimates of the necessary atmospheric parameters, Bates and Massey (1946), using Chapman's formula (1), found that the location of the F_1 layer appeared to be in fair agreement with what it would be if the layer were produced by the ionization of atomic oxygen,

$$O + h\nu \rightarrow O^+ + e. \qquad \ldots\ldots(3)$$

They also drew attention to the fact that the flux of photons from the conventional 6,000° K. black body model of the Sun is of the order required to account for the observed rate of ionization. It seemed plausible therefore to assume that process (3) is responsible for the F_1 layer ; and, as will be recalled, the F_2 layer does not require a different mechanism. The possibility that the ionization of atomic nitrogen,

$$N + h\nu \rightarrow N^+ + e, \qquad \ldots\ldots(4)$$

gives some contribution to both layers is not, of course, excluded.

As far as the E layer is concerned, Bates and Massey (1946) pointed out that if the theory just mentioned were accepted there would be a severe restriction on the processes that could be involved as radiation of quantal energy between 13·5 ev. (the first ionization potential of O) and well over 25 ev. could not penetrate sufficiently deeply into the atmosphere. Two suggestions have been made: Nicolet (1945) has proposed that the responsible mechanism is pre-ionization of molecular oxygen in the energy range 12·2 to 13·5 ev., and Hoyle and Bates (1948) have given consideration to the possibility of general ionization by photons of very high energy (perhaps about 325 ev.) originating in the solar corona. Both must be regarded as tentative as they depend respectively on the unproven assumptions that pre-ionization of molecular oxygen can occur and that the coronal emission is of sufficient intensity. The observational evidence is still indecisive, but tends to favour the former hypothesis rather than the latter (cf. Bates 1949 a).

---

* Certain reservations on this will be found in the original paper.

It is clear that information on the absorption cross sections of atomic oxygen and nitrogen finds wide application in the detailed discussion of layer formation. In view of this Bates and Seaton (1950) have recently attempted to refine the earlier calculations by including allowance for the effects of electron exchange and by making use of both the dipole moment and dipole velocity formulae for the transition matrix element. The new cross sections obtained are given in the Table. They are appreciably below (actually about half) the old values. While the corrections are not unduly great, their sense, unfortunately, is such as to increase some minor discrepancies already appearing in the provisional theory outlined above. Consequently it is desirable to reconsider the position.

	Ionization potential (ev.)	Absorption cross section at spectral head (A) (10^{-18} cm^2)
Atomic oxygen (O)	13·5	2·6
Atomic nitrogen (N)	14·5	9

The simplest method of studying the significance of the cross sections is to use them in conjunction with formula (1) to determine the atmospheric parameter $Hn(\mathrm{X})$. For the sake of definiteness it is convenient to adopt a specific value of $\cos\chi$; 0·8 is a sufficiently representative figure.

Consider the ionization of atomic oxygen by photons in the energy range 13·5 to 16·9 ev. As can be seen at once[*] from the Table, $Hn(\mathrm{O})$ is about $3 \times 10^{17}/\mathrm{cm}^2$. Now at the level of the F_1 layer H and $n(\mathrm{O})$ are usually assumed to be 30 km. and $3 \times 10^{10}/\mathrm{cm}^3$, respectively[†], giving a value of $Hn(\mathrm{O})$ equal to $1 \times 10^{17}/\mathrm{cm}^2$, which is three times smaller than the theory demands. It is difficult to decide whether or not the discrepancy is real. The accuracy of the atmospheric data is very uncertain[‡], and it may well be that the conflict is caused by the combination of errors in H and $n(\mathrm{O})$, and perhaps to some extent from a decrease of the recombination coefficient with altitude (which would tend to raise the maximum of the layer). But to obtain agreement it is necessary to postulate *greater* values of the first two quantities than are normally assumed. A change of this nature in the supposed structure of the atmosphere, while not inconceivable, cannot be accepted lightly, involving as it does an increase in the already high temperature. It is worth considering therefore whether there is any possibility that atomic oxygen is *not* the constituent responsible for the F_1 layer.

[*] Molecular nitrogen is so transparent in the spectral region concerned that no correction need be made for absorption by it. We are grateful to Dr. W. C. Price and Dr. R. E. Worley for information in this connection.

The adoption of the A appropriate to the spectral head is only justified if the flux of the solar radiation decreases rapidly with energy (as is the case for a 6,000° K. black body). If the fall off proves to be slow, a higher mean A should be taken to allow for ionization to $\mathrm{O}^+(^2\mathrm{D})$ and $\mathrm{O}^+(^2\mathrm{P})$. The effect of this would be to decrease the discrepancy mentioned in the text.

[†] The data used here (and elsewhere in the paper) are taken from the 'model atmosphere' proposed by Bates and Massey (1946). The basis for the estimates is further discussed in a recent review (Bates 1949 b).

[‡] Harang (1945), for example, advocates scale heights and densities that are rather different from those that have been adopted.

On account of the intensity of the ionization likely to result from process (3) it is reasonable to suppose that its maximum is located in some layer. As an alternative to the F_1 layer the E layer might be suggested. In it H is about 10 km. and $n(O)$ is about $2 \times 10^{12}/\text{cm}^3$, so that $Hn(O)$ is approximately $2 \times 10^{18}/\text{cm}^2$. The theoretical figure given in the previous paragraph must not be compared directly with this. It is necessary to make allowance for the probable rapid upward increase of the scale height in the region concerned. Assuming for simplicity that it is a linear function of the altitude h, Chapman's formula is replaced by

$$n(X) = (1 + dH/dh)\cos\chi/HA, \qquad \ldots\ldots(5)$$

the H in the denominator being that at the layer maximum. dH/dh may be as great as unity, so that the value of $Hn(O)$ based on the calculated cross section now becomes perhaps $6 \times 10^{17}/\text{cm}^2$. Though the discrepancy is thus no larger than for the F_1 layer it is rather less easy to ignore as H and $n(O)$ are better determined in the E layer; further, H can scarcely be below the 10 km. adopted so that the entire error would have to be attributed to $n(O)$, which would hence need to be $6 \times 10^{11}/\text{cm}^3$ instead of the estimated $2 \times 10^{12}/\text{cm}^3$. The E_2 layer, being slightly above the normal E (or E_1) layer, is clearly more satisfactory as far as the location requirement is concerned. However, in view of the fact that it can only be observed occasionally, the possibility that an ionizing action as effective as that under discussion is involved must be treated with caution. Radio evidence on the factors governing its appearance, and on the density of ionization in the region between the two main layers, would be useful.

To complete our exploratory enquiry we must examine whether any process other than (3) is capable of producing the F_1 and F_2 layers: for unless there is some such process the various suggestions put forward for consideration must be abandoned. The ionization of atomic nitrogen at once presents itself as a possible mechanism. Taking H to be 30 km. (as before) the value of $n(N)$ that is required in the Chapman-like F_1 layer can be seen to be $3 \times 10^{10}/\text{cm}^3$ (compared with a total particle concentration of $1 \times 10^{11}/\text{cm}^3$). Though this figure would imply a high degree of dissociation it cannot be said to be impossible (cf. Bates 1949 b). The integrated rate of ionization/cm² column in the F layers is, of course, several times greater than that in the E layers. It might be thought at first that this is in serious conflict with the fact that the ionization potential of N is higher than that of O (14·5 ev. compared with 13·5 ev.). If the 6,000° K. black body model of the Sun is used the difficulty certainly appears very acute. However, such a model, though providing a useful general guide for some purposes, does not necessarily correspond with reality; and indeed recent work by Hoyle* (1949) suggests that the intensity of the solar radiation beyond the Lyman limit falls off so slowly with increasing frequency that the photon flux above 14·5 ev. is actually greater than that between 13·5 and 14·5 ev.

2.2. In summarizing the position it is convenient to regard the development of the theory of the formation of the ionized layers as progressing in two main stages: (*a*) the listing of all processes that might conceivably have to be taken into consideration, and the investigation of which combinations of them might give a plausible account of the observed ionospheric structure; and (*b*) the detailed quantitative examination of the alternatives to decide which is correct. It seems probable that the speculative but essential first stage is largely completed. However, though the conditions that must be satisfied are fully appreciated, and though some important restrictions on the possibilities are now realized, much of the second and far more difficult stage has still to be traversed. In order to discriminate unambiguously between the theories suggested it is necessary to have precise information on certain atmospheric and other parameters. The task of attaining the required degree of accuracy is formidable. Until it is accomplished the exclusive advocacy of a particular theory is unjustified. In particular no firm decision can yet be reached as to whether the F layers arise mainly from atomic oxygen, or whether the maximum of the ionization rate due to this constituent occurs at a lower level and the principal active constituent responsible for the F layers is atomic nitrogen.

---

* We wish to thank Mr. Hoyle for showing us his results before publication and for many stimulating discussions.

§ 3. THE D LAYER

3.1. The theories of the origin of the D layer cannot be properly discussed without some knowledge of the rate of ionization to be explained. Elaborate calculations are not merited at present and we will therefore make estimates of this rate only at two specific levels, the 75 km. and 90 km. levels. As these are believed to be near the base and near the maximum of the layer, respectively, the salient features will be illustrated adequately. It must be stressed that the figures we deduce are essentially of a preliminary nature and will doubtless have to be revised later. Our intention in giving them is merely to provide an indication of the problem that is presented.

Radio scientists have published very little quantitative data on D layer ionization. However Mr. W. R. Piggott has been generous enough to acquaint us with the results he has obtained during the past few years from absorption measurements. It appears probable from his investigations that at the 90 km. level the electron concentration, $n(\mathrm{e})$, is about $1{\cdot}5 \times 10^4/\mathrm{cm}^3$ at mid-day, and the effective recombination coefficient α is about $3 \times 10^{-8}\,\mathrm{cm}^3/\mathrm{sec}$. In the near equilibrium conditions prevailing the effective rate of electron production, q, may be taken as equal to $\alpha n(\mathrm{e})^2$. From the figures just quoted its value can be seen to be approximately $7/\mathrm{cm}^3/\mathrm{sec}$.

The lower part of the layer is best explored by the study of the reflection of long radio waves. Dr. K. Weekes has informed us that the most recent work using this technique indicates that $n(\mathrm{e})$ is some $2{\cdot}5 \times 10^2/\mathrm{cm}^3$ at or near the 75 km. level at mid-day. While α in this region has not as yet been determined, some provisional information on it can be obtained: thus by using the theoretical ionization curves compiled by Chapman (1931 b) it can readily be shown that in order that there be sufficient time for $n(\mathrm{e})$ to reach the observed value, q at noon must be at least some $6 \times 10^{-2}/\mathrm{cm}^3/\mathrm{sec}$., which implies that α must be at least of order $1 \times 10^{-6}/\mathrm{cm}^3/\mathrm{sec}$.

Now α and q are not themselves of fundamental significance: they are simply parameters appearing in the familiar equation governing the variation of $n(\mathrm{e})$ with time:

$$dn(\mathrm{e})/dt = q - \alpha n(\mathrm{e})^2. \qquad \ldots\ldots(6)$$

For our purpose we are interested in the actual electronic and ionic recombination coefficients, α_e and α_i respectively, and in the true rate of photo-ionization, $\mu n(\mathrm{X})$.

The two sets of quantities are related. Introducing λ to represent the ratio of th concentration of negative ions to that of electrons it may be proved that

$$\alpha \simeq \alpha_e + \lambda \alpha_i \qquad \ldots\ldots(7)$$

and

$$\mu n(X) = (1+\lambda)q \qquad \ldots\ldots(8)$$

(Appleton and Sayers 1938, Bates and Massey 1946). Formula (8) is of particular importance in showing that q is less than $\mu n(X)$. The figures given in the previous paragraphs must therefore only be regarded as setting a lower limit to the photo-ionization rate.

An estimate of λ can be made by studying the equilibrium between the reactions leading to the formation and destruction of negative ions. On account of the relatively high gas densities involved it is probable that electrons become attached to oxygen molecules mainly through the Block–Bradbury process (Bates and Massey 1946). The coefficient η associated with this depends on the concentration of molecular oxygen: analysis of the scanty experimental data available indicates that its value may be some $2{\cdot}5 \times 10^{-14}\,\mathrm{cm^3/sec.}$ at 75 km. (where $n(O_2)$ is taken as $4 \times 10^{14}/\mathrm{cm^3}$) and some $1 \times 10^{-14}\,\mathrm{cm^3/sec.}$ at 90 km. (where $n(O_2)$ is taken as $8 \times 10^{13}/\mathrm{cm^3}$). During the day the O_2^- ions formed must suffer photo-detachment,

$$O_2^- + h\nu \rightarrow O_2 + e. \qquad \ldots\ldots(9)$$

Crude calculations suggest that the rate ρ at which this occurs may be perhaps about 0·5 per negative ion per second. Hence from the formula

$$\lambda \simeq \eta n(O_2)/\rho \qquad \ldots\ldots(10)$$

it can be seen that

$$\lambda_{75\ \mathrm{km.}} \simeq 20 \qquad \ldots\ldots(11)$$

$$\lambda_{90\ \mathrm{km.}} \simeq 2.^* \qquad \ldots\ldots(12)$$

Unfortunately the uncertainties in η and ρ are so serious that these figures should only be regarded as indicating the possible order of magnitude at each of the two levels concerned. Equation (7) however provides at least some confirmation that $\lambda_{75\ \mathrm{km.}}$ is large, for otherwise it is difficult to explain an α of $10^{-6}\,\mathrm{cm^3/sec.}$ since neither α_e nor α_i are likely to be much in excess of $10^{-8}\,\mathrm{cm^3/sec.}$ (Bates and Massey 1946, 1947).

Finally substitution in the expression for the photo-ionization rate (8) yields

$$\mu n(X)_{75\ \mathrm{km.}} \simeq 1{\cdot}3/\mathrm{cm^3/sec.}, \qquad \ldots\ldots(13)$$

$$\mu n(X)_{90\ \mathrm{km.}} \simeq 20/\mathrm{cm^3/sec.} \qquad \ldots\ldots(14)$$

The fact that these estimates may be in error by a considerable factor must be kept constantly in mind throughout the remainder of the discussion.

3.2. A number of theories of D layer ionization (normal and enhanced) have been put forward but have not been developed sufficiently quantitatively for their merits and defects to be properly assessed. For reference purposes we will first summarize them briefly.

---

* It may be noted that below the 75 km. level λ increases rapidly downwards, probably indeed as the square of the gas density. The effective recombination coefficient α must increase at least at the same rate (or faster if three body processes contribute to α_i) and must attain a very large value in the region immediately beneath the base of the normal layer.

In 1936 Chapman and Price drew attention to the close coincidence that exists between λ1215·6 A., the wavelength of the L(α) line of hydrogen, and λ1217·6 A., the wavelength of the radiation required to raise atomic oxygen from the metastable state, $(p^4)^1S$, to the highly excited state, $(p^3s)^1P$; they also pointed out that an atom in this latter state might suffer ionization as a result of a collision of the second kind. Remarking that the solar spectrum may contain L(α) as a broad emission line they suggested that appreciable atmospheric ionization might be produced by the sequence

$$O(^1S) + L(\alpha) \rightarrow O(^1P), \qquad \ldots\ldots(15)$$

$$O(^1P) + X \rightarrow O^+(^4S) + X + e + 0{\cdot}75\ \text{ev}. \qquad \ldots\ldots(16)$$

Later Martyn, Munroe, Higgs and Williams (1937) proposed that this mechanism might be responsible for the D layer.* Other authors have favoured ordinary photo-ionization processes (Mitra, Bhar and Ghosh 1939, Nicolet 1945, Bates and Massey 1946). The more obvious possibilities are:

$$O_2 + h\nu \rightarrow O_2^+ + e, \qquad \ldots\ldots(17)$$

$$O_3 + h\nu \rightarrow O_3^+ + e, \qquad \ldots\ldots(18)$$

$$NO + h\nu \rightarrow NO^+ + e. \qquad \ldots\ldots(19)$$

These possibilities have one important characteristic in common—they involve radiation of comparatively high photon energy.† Jouaust and Vassy (1941) and later Vassy and Vassy (1942) have claimed that laboratory experiments indicate that such radiation (and indeed radiation of photon energy down to about 7ev.) is so strongly absorbed by air that it is impossible for it to penetrate to the required levels. To avoid this apparent difficulty they suggested that the mechanism operative might be the photo-ionization of atomic sodium,

$$Na + h\nu \rightarrow Na^+ + e, \qquad \ldots\ldots(20)$$

the threshold energy for which is only 5·1 ev.

To differentiate between the various theories outlined it is necessary to examine each critically. We will begin by considering the sodium theory, as in view of the remarks in the previous paragraph it might seem to be the only one tenable. $\mu n(\text{Na})$, the contribution to the total ionization rate from (20), can be calculated at least approximately. It depends on $n(\text{Na})$, the concentration of atomic sodium, A, the absorption cross section associated with the process, and S, the incident photon flux in the relevant energy range. Information on all three is available.

(i) Observations on the twilight enhancement of the D lines enable $n(\text{Na})$ to be determined. The most recent work (that of Barbier (1948)) gives a value of approximately $3 \times 10^3/\text{cm}^3$ at 75 km. and $6 \times 10^2/\text{cm}^3$ at 90 km. Variations of course occur—concentrations much smaller than those quoted have been reported.

---

* The original theory referred to the fade-out enhancement, but for completeness we will consider it extended to include the normal ionization.

† L(α) corresponds to 10·1 ev. The ionization potentials of O_2 and NO are known to be 12·2 and 9·1 ev.; that of O_3 has not been measured, but the theoretical work of Mulliken (1942) suggests it to be 12·5 ev. (which is consistent with the estimated value obtained from the standard dissociation–ionization cycle). It may be noted that the ionization potentials of the other likely constituents are also high (H_2O, 12·6 ev.; N_2O, 12·7 ev.; CO_2, 13·7 ev.; CO, 13·9 ev.; H_2, 15·4 ev.).

(ii) Recent measurements on A (Ditchburn and Jutsum, unpublished *) show that at the ionization limit A equals $1{\cdot}16 \times 10^{-19}\,\mathrm{cm}^2$; immediately above it falls off rapidly with increasing frequency.

(iii) By means of rocket-borne instruments the solar spectrum in the main region of importance (λ2415 A. to λ2200 A.) has now been surveyed. Preliminary curves showing the intensity distribution are given in the reports of the U.S. Naval Research Laboratory (1948). The mean absolute value is about one-tenth that from a 6,000° K. black body.

Combining the data from these sources it can readily be shown that

$$\mu n(\mathrm{Na})_{75\,\mathrm{km.}} \simeq 3 \times 10^{-2}/\mathrm{cm}^3/\mathrm{sec.}\dagger \qquad \ldots\ldots(21)$$

and

$$\mu n(\mathrm{Na})_{90\,\mathrm{km.}} \simeq 5 \times 10^{-3}/\mathrm{cm}^3/\mathrm{sec.}\dagger \qquad \ldots\ldots(22)$$

These rates are very small compared with the rates apparently required (§ 3.1). Though (20) is probably augmented by similar processes involving other metallic elements (traces of which are presumably present in the upper atmosphere) it would seem most unlikely that the effect could be sufficient to make the total contribution appreciable except perhaps at the extreme base of the layer. A further discrepancy also arises. Since the active radiation is not appreciably attenuated, the ionization rate during the day would be expected to be essentially independent of the solar zenith angle. This expectation is in direct conflict with the observations of radio workers. While these are consistent with the existence of a small component which remains constant during the day (cf. Gockel 1948) they show unmistakably that the main ionization rate is closely controlled by the elevation of the Sun (Budden, Ratcliffe and Wilkes 1939). It would seem then that some additional mechanism must be sought.

Before proceeding further we must reconsider the claim of Jouaust, Vassy and Vassy that only low energy photons can penetrate to the level of the D layer. This was based mainly on measurements made by Schneider (1940) on the absorption by air in the extreme ultra-violet (λ380 A. to λ1600 A.). These certainly indicate that photons of energy greater than about 12 ev. are very effectively absorbed. Hopfield (1946) has recently published evidence confirming this. His work is particularly valuable as (unlike Schneider) he used a *continuous* light source. The spectrograms he gives show that the radiation, in the limited range under discussion, is almost completely extinguished in passing through a 0·4 cm. column of air at S.T.P. Now the equivalent path length down to the 90 km. level is about 10 cm. Hence, as can readily be verified, the residual intensity even here is at most only some 10^{-11} times the incident intensity. It is apparent therefore that processes (17) and (18) cannot yield appreciable ionization: the necessary detailed calculations are elementary and need scarcely be presented.

The situation with regard to photons of energy below 12 ev. is less simple. Jouaust, Vassy and Vassy adopted, without reservation, the absorption cross sections tabulated by Schneider. However the figures given are not uniformly reliable. Schneider himself states that those relating to the more transparent spectral regions should be regarded only as approximate *upper limits*. One of these regions is of special interest as it includes L(α). Williams (1940) and Preston (1940) have independently studied the absorption of this individual line. Their results did not agree—a large cross section was obtained by Williams and a small cross section by Preston. Jouaust, Vassy and Vassy argued that the former result should be accepted (with minor modifications) and that the latter should be rejected; the accord this gave with their interpretation of Schneider's work seemed reasonable. We have however been in communication with Dr. Williams who has informed us that he has performed further measurements and now considers that Preston's result is correct—that is, that the absorption cross section of L(α) in air is only about $2 \times 10^{-21}\,\mathrm{cm}^2$. The extremely low attenuation rate this implies is of great significance. It means that the line penetrates deep into the atmosphere: to reduce the intensity by a factor of 10 an equivalent path of some 42 cm. must be traversed. Doubtless rocket experiments will soon provide conclusive evidence on the matter.

---

* We wish to thank Professor R. W. Ditchburn for placing at our disposal these results obtained in his laboratory.

† Rates previously estimated (Bates 1947) were greater owing to the adoption of the absorption cross section computed by Rudkjöbing (1940), which is rather larger than that of Ditchburn and Jutsum, and to the acceptance of the view then widely held that the solar emission in the spectral region concerned was somewhat in excess of that from a black body at 6,000° K. The values originally given for the ratio [Na$^+$]/[Na] are thus too great, but even with the new data the degree of ionization of atmospheric sodium remains remarkably high.

Hopfield's investigation gives no information regarding L(α) as atomic hydrogen was present in the apparatus used. It does however indicate that air has several other 'windows' of high transparency in the photon energy range between 10 and 11 ev. As far as can be judged from the spectrograms published the radiation in these is not seriously attenuated in a path 4·2 cm. long. This suggests that the absorption cross section is not more than some $10^{-20}\,\mathrm{cm}^2$ (perhaps about $5 \times 10^{-21}\,\mathrm{cm}^2$).

We are now in a position to discuss the two remaining theories of the origin of the D layer.

In estimating the ionization rate $\mu n(\mathrm{O})$ resulting from (15) and (16) it is convenient to express the L(α) intensity emanating from the Sun as x times the intensity of the continuum (at the corresponding frequency ν) emitted by a black body at a temperature T. Before reaching the Earth the radiation is diluted by a factor f, and before penetrating to the D layer it is attenuated by a further factor g. If ω_A and ω_B are the statistical weights of the initial state A, and final state B, of the oxygen atom, if r is the ratio of the number of spontaneous transitions from B to A to the total number from B to any state, and if ζ is the rate coefficient associated with the ionizing collisions postulated, then it can be seen that

$$\mu n(\mathrm{O}) = \left\{\frac{\omega_\mathrm{B}}{\omega_\mathrm{A}} xfg \exp(-\boldsymbol{h}\nu/\boldsymbol{k}T)\right\} r\zeta n(\mathrm{X}) n(\mathrm{O}^1\mathrm{S}), \qquad \ldots\ldots(23)$$

$\boldsymbol{k}$ being Boltzmann's constant. The values of all the quantities except x appearing in this formula are either known or can be limited in some way.

We will first summarize briefly the information available, and then consider what are the demands on x made by the theory. The values of ω_A and ω_B are 1 and 3 respectively. Because of the presence of the adjustable factor x we may assign to T the conventional value 6,000° K. f is $5{\cdot}4 \times 10^{-6}$ (a standard figure). From Preston's measurements on absorption it can be shown that $g_{75\,\mathrm{km.}}$ is of order 6×10^{-2} and that $g_{90\,\mathrm{km.}}$ is approximately 5×10^{-1}. Estimates of the relative transition probabilities suggest that r is about 0·3 (and in any event it must be less than unity). Since the ionizing collisions violate the spin conservation rule their efficiency would not be expected to be high and it therefore seems improbable that ζ is greater than $1 \times 10^{-11}\,\mathrm{cm}^3/\mathrm{sec}$. X, the constituent involved in the ionizing collisions, is unidentified, but a sufficient upper limit to $n(\mathrm{X})$ is provided by N, the total particle concentration, which at 75 km. is $2 \times 10^{15}/\text{cm}^3$ and at 90 km. is $4 \times 10^{14}/\text{cm}^3$. A tentative estimation of $n(\text{O}^1\text{S})$ can be obtained by considering the mechanisms leading to the formation and destruction of metastable oxygen atoms. Bates and Massey (1946) have studied the equilibrium between the processes

$$\text{O}(^3\text{P}) + h\nu \rightleftharpoons \text{O}(^1\text{S}) \qquad \ldots\ldots(24)$$

and have shown that it gives

$$n(\text{O}^1\text{S}) \simeq 3 \times 10^{-11} n(\text{O}^3\text{P}). \qquad \ldots\ldots(25)$$

Detailed calculations have recently been made by Penndorf (1949) on the amount of atomic oxygen in the upper atmosphere. These indicate that $n(\text{O}^3\text{P}_{90\ \text{km.}})$ is unlikely to exceed about $10^{13}/\text{cm}^3$ (and it may well be much smaller). $n(\text{O}^3\text{P}_{75\ \text{km.}})$ was not computed but it must clearly be less than the figure just quoted; $4 \times 10^{12}/\text{cm}^3$ seems a reasonable upper limit to take. Hence from (25) $n(\text{O}^1\text{S}_{75\ \text{km.}})$ is at most $120/\text{cm}^3$ and $n(\text{O}^1\text{S}_{90\ \text{km.}})$ is at most $300/\text{cm}^3$.* Finally, numerical substitution in (23) yields

$$\mu n(\text{O})_{75\ \text{km.}} < 2 \times 10^{-9} x/\text{cm}^3/\text{sec.} \qquad \ldots\ldots(26)$$

and

$$\mu n(\text{O})_{90\ \text{km.}} < 1 \times 10^{-8} x/\text{cm}^3/\text{sec.} \qquad \ldots\ldots(27)$$

In spite of the uncertainty with regard to x, these ionization rates appear so slow that the possibility that processes (15) and (16) might account for the ordinary D layer would seem to be excluded. Many authors consider that the intensity of $L(\alpha)$ becomes extremely high during the chromospheric eruptions with which radio fade-outs are associated and that x may actually be of order 10^6 for a major disturbance (Hunter 1942–43). A still greater intensity is required to explain the enhancement of the D layer, so that even for this special phenomenon the theory shows little promise.

It is not unlikely that nitric oxide is an important constituent of the upper atmosphere. Price (1942–43), for example, has pointed out that it is one of the products of dissociation of nitrous oxide which is probably present in considerable concentrations (Sutherland and Callendar 1942–3). There is some reason too for believing that nitrogen as well as oxygen occurs partly in the free atomic form (Bates 1949 b); if this is the case it would be expected that nitric oxide would be formed by chemical action. Again Durand, Oberly and Tousey (1949) have recently reported what may prove to be direct confirmatory evidence. In the course of the analysis of a spectrum of the Sun taken during a rocket flight to an altitude of 55 km. they discovered a broad band lying between $\lambda 2200$ A. and $\lambda 2300$ A.; they suggested that it may be caused by absorption by atmospheric nitric oxide. If this interpretation is correct the number of NO molecules above the 55 km. level must be of order $10^{18}/\text{cm}^2$ column. Nothing can be said at present about the distribution, but it is at least possible that in the region concerned it is much the same as that of the main gases.

Lack of basic data makes it impossible to investigate the photo-ionization of nitric oxide sufficiently quantitatively to establish whether or not it is responsible for the formation of the D layer, but in view of the unsatisfactory position of the alternative theories considered the possibility appears to be at least worth examining. A calculation so speculative does not justify detailed presentation, and we state briefly the assumptions made and the conclusions reached: (i) Consistent with a total NO content of $10^{18}/\text{cm}^2$ column we take $n(\text{NO})_{75\text{ km.}}$ to be $2 \times 10^{11}/\text{cm}^3$ and $n(\text{NO})_{90\text{ km.}}$ to be $4 \times 10^{10}/\text{cm}^3$. (ii) It is known that the cross section associated with process (19) is low (Price 1942–43): we arbitrarily adopt a value of $5 \times 10^{-20}\,\text{cm}^2$. (iii) In the absence of more exact information we regard the quiescent Sun as a 6,000° K. black body. Accepting Preston's (1940) and Hopfield's (1946) work we treat the atmosphere as opaque except for two groups of windows: in one (containing $L(\alpha)$ and of width perhaps 5 A.) the absorption cross section is about $2 \times 10^{-21}\,\text{cm}^2$, in the other (of width perhaps 20 A.) the absorption cross section is $5 \times 10^{-21}\,\text{cm}^2$. This gives the photon flux at 75 km. as about $1 \times 10^{9}/\text{cm}^2/\text{sec.}$ and that at 90 km. about $2 \times 10^{10}/\text{cm}^2/\text{sec.}$

---

* We should perhaps mention that Chapman's association–excitation process $3\text{O} \rightarrow \text{O}_2 + \text{O}(^1\text{S})$ does not lead to greater concentrations than those given. This can readily be demonstrated by a simple argument based on the known rate of emission of $\lambda 5577$ A. during the night (when the amount of atomic oxygen is not sensibly less than during the day (Penndorf 1949)). The neglected de-activation collisions will, of course, tend to reduce $n(\text{O}^1\text{S})$.

Combining (i), (ii) and (iii) it can be seen at once that

$$\mu n(\text{NO})_{75\text{ km.}} \simeq 10/\text{cm}^3/\text{sec.}$$

$$\mu n(\text{NO})_{90\text{ km.}} \simeq 40/\text{cm}^3/\text{sec.}$$

These rates are considerable, and it may be that the photo-ionization of nitric oxide actually does yield the major part of the ordinary D layer. As need scarcely be emphasized, the estimates made are too crude to justify an assured statement.

The nitric oxide theory appears quite attractive as far as the enhancement during fade-outs is concerned, for if $L(\alpha)$ is greatly intensified a high ionization rate would certainly occur unless the values given for $n(\text{NO})$ are grossly in excess. It is interesting to note that on account of the difference in the transparency of the two groups of 'windows' in the atmosphere an intensification of $L(\alpha)$ would tend to cause a lowering of the layer. Radio measurements give some evidence of such an effect (Budden, Ratcliffe and Wilkes 1939).

Mention must, however, be made of a possible difficulty associated with any purely photo-ionization theory of fade-outs. Suppose that the height of reflection of long radio waves is normally h_N, and that is is reduced to h_F during a fade-out. τ, the time for recovery to $\frac{1}{2}(h_N + h_F)$, is simply the time taken for the electron concentration at this mean level to fall from its enhanced value, $n_F(e)$, to the value required for reflection, $n_R(e)$, and is therefore given by

$$\tau = \frac{1 - n_R(e)/n_F(e)}{\alpha n_R(e)}. \qquad \ldots\ldots(28)$$

Now the experiments of Bracewell and Straker (1949) have shown that τ is not more than some 150 sec.; $n_R(e)$ is about $250/\text{cm}^3$; $n_F(e)$ is unknown, but provided it is not too small the numerator of the fraction in (28) can be taken as approximately unity. Hence it would seem that α is at least of order $3 \times 10^{-5}\,\text{cm}^3/\text{sec.}$ at the mean level (which may be a few kilometres below the normal height of reflection). Such a coefficient, though perhaps not impossible to accept,* is certainly large enough to cause some uneasiness. The obvious alternative to recombination is that the recovery is due to the attachment of electrons to neutral particles to form negative ions. This process could clearly proceed at a sufficient rate, but it requires that during the flare the ratio of the concentration of negative ions to that of electrons should be reduced: and as far as can be judged ordinary photodetachment would only yield a reduction if there were an appreciable increase in the *total* energy flux from the Sun—an event that seems inconceivable.

---

* For example the estimate of λ made in § 3.1 may be too low. See also footnote on page 134.

ACKNOWLEDGMENT

One of us (M.J.S.) wishes to express his thanks to the Ministry of Education for the award of a grant.

REFERENCES

ALLEN, C. W., 1948, *Terr. Magn. Atmos. Elect.*, **53,** 433.

APPLETON, E. V., and SAYERS, J., 1938, *Union Radio Scient. Inst.*, **78,** 272.

BARBIER, D., 1948, *Ann. Geophys.*, **4,** 193.

BATES, D. R., 1939, *Mon. Not. Roy. Astr. Sc.*, **100,** 25; 1947, *Terr. Magn. Atmos. Elect.*, **52,** 71; 1949 a, *Proc. Roy. Soc.* A, **196,** 562; 1949 b, *Mon. Not. Roy. Astr. Soc.* **109,** 216.

BATES, D. R., BUCKINGHAM, R. A., MASSEY, H. S. W., and UNWIN, J. J., 1939, *Proc. Roy. Soc.* A, **170,** 322.

BATES, D. R., and MASSEY, H. S. W., 1946, *Proc. Roy. Soc.* A, **187,** 261; 1947, *Ibid.*, **192,** 1.

BATES, D. R., and SEATON, M. J., 1950, *Mon. Not. Roy. Astr. Soc.* (in the press).

BRACEWELL, R. N., and STRAKER, T. W., 1949, *Mon. Not. Roy. Astr. Soc.*, **109,** 28.

BRADBURY, N. E., 1938, *Terr. Magn. Atmos. Elect.*, **43,** 55.

BUDDEN, K. G., RATCLIFFE, J. A., and WILKES, M. V., 1939, *Proc. Roy. Soc.* A, **171,** 188.

CHAPMAN, S., 1931 a, *Proc. Roy. Soc.* A, **132,** 353; 1931 b, *Proc. Phys. Soc.*, **43,** 26, 433.

CHAPMAN, S., and PRICE, W. C., 1936, *Rep. Prog. Phys.*, **3,** 55 (London: The Physical Society).

DUFAY, M., and DUFAY, J., 1947, *C.R. Acad. Sci., Paris*, **224,** 1834.

DURAND, E., OBERLY, J. J., and TOUSEY, R., 1949, *Astrophys. J.*, **109,** 1.

GOCKEL, H., 1948, *Ann. Geophys.*, **4,** 232.

HARANG, L., 1945, *Geofys. Publ., Oslo*, **16,** No. 6.

HOPFIELD, J. J., 1946, *Astrophys. J.*, **104,** 208.

HOYLE, F., 1949, *Recent Advances in Solar Physics* (Cambridge: University Press).

HOYLE, F., and BATES, D. R., 1948, *Terr. Magn. Atmos. Elect.*, **53,** 51.

HUNTER, A., 1942–43, *Rep. Prog. Phys.*, **9,** 5 (London: The Physical Society).

JOUAUST, R., and VASSY, E., 1941, *C.R. Acad. Sci., Paris*, **213,** 139.

MARTYN, D. F., MUNRO, G. H., HIGGS, A. J., and WILLIAMS, S. E., 1937, *Nature, Lond.*, **140,** 603.

MITRA, S. K., BHAR, J. N., and GHOSH, S. P., 1939, *Indian J. Phys.*, **12,** 455.

MULLIKEN, R. S., 1942, *Rev. Mod. Phys.*, **14,** 204.

NICOLET, M., 1945, *Mem. R. Met. Inst., Belgium.*, **19,** 1.

PENNDORF, R., 1949, *J. Geophys. Res.*, **54,** 7.

PRESTON, W. M., 1940, *Phys. Rev.*, **57,** 887.

PRICE, W. C., 1942–43, *Rep. Prog. Phys.*, **9,** 10 (London: The Physical Society).

RUDKJÖBING, A., 1940, *Publ. Kbh. Obs.*, **18,** 1.

SCHNEIDER, E. G., 1940, *J. Opt. Soc. Amer.*, **30,** 128.

SUTHERLAND, G. B. B. M., and CALLENDAR, G. S., 1942–43, *Rep. Prog. Phys.*, **9,** 18 (London: The Physical Society).

U.S. Naval Research Lab. Report, 1948, No. R3120.

VASSY, A., and VASSY, E., 1942, *Cahiers de Physique*, **9,** 28.

WILLIAMS, S. E., 1940, *Nature, Lond.*, **145,** 68.

PHYSICAL SOCIETY CONFERENCE AT CAMBRIDGE

14th to 16th JULY 1949

Thursday, 14th July.

2.30 – 5.15 p.m.

Session 1.—*The regular behaviour of long and very long waves returned from the ionosphere.*

Speakers : J. A. Ratcliffe, *Cavendish Laboratory.*
K. W. Tremellen, *Marconi's Wireless Telegraph Co. Ltd.*
W. T. Sanderson, *Decca Navigator Co. Ltd.*
C. Williams, *Royal Aircraft Establishment.*
R. N. Bracewell, *Cavendish Laboratory.*

Friday, 15th July.

10 a.m. – 12.30 p.m.

Session 2.—*The regular behaviour of medium and short waves.*

Speakers : W. J. G. Beynon, *University College of Swansea.*
A. F. Wilkins, *Radio Research Station, Slough.*
A. B. Whatman, *Ministry of Supply.*

2.30 – 4.0 p.m.

Session 3.—*The irregular behaviour associated with solar events.*

Speakers : W. R. Piggott, *Dept. of Scientific and Industrial Research.*
K. Weekes, *Cavendish Laboratory.*
R. Rivault, *Université de Poitiers*, will read a paper by R. Bureau of the *Laboratoire National de Radioélectricit, Bagneux.*

4.30 – 6.0 p.m.

Session 4.—*The formation of ionized regions.*

Speakers : K. Weekes, *Cavendish Laboratory.*
D. R. Bates, *University College, London.*

Saturday, 16th July.

10 a.m. – 12.30 p.m.

Session 5.—*Irregularities in the horizontal plane in the ionosphere.*

Speakers : J. W. Findlay, *Cavendish Laboratory.*
A. C. B. Lovell, *Manchester University.*
G. Millington, *Marconi's Wireless Telegraph Co. Ltd.*
W. Ross, *Radio Research Station, Slough*
R. Rivault, *Université de Poitiers.*

ABSTRACTS OF PAPERS

SESSION 1

THE REGULAR BEHAVIOUR OF LONG AND VERY LONG WAVES RETURNED FROM THE IONOSPHERE

The Regular Behaviour of Long and Very Long Waves Returned from the Ionosphere

By J. A. Ratcliffe
Cavendish Laboratory, Cambridge

These remarks will (*a*) define the scope of this session, (*b*) indicate how the main contributions fit into the plan of the discussion, and (*c*) outline some of the work carried out in Cambridge.

The frequencies considered lie between 10 kc/s. and 300 kc/s. Observations have mainly been near the frequencies of 16 kc/s. (GBR Rugby) and 100 kc/s. (Decca). Academic workers have generally concentrated on ionospheric reflection near vertical incidence and commercial organizations on reflection at oblique incidence. This session is concerned only with regular behaviour of the ionosphere.

By day the wave reflected at vertical incidence on 100 kc/s. is very weak (reflection coefficient $\rho=0{\cdot}003$ in summer), but on 16 kc/s. is quite strong ($\rho=0{\cdot}15$ in summer), and there appears to be a difference in the mechanism of reflection. Experiments made at Cambridge on other frequencies to find how this reflection coefficient changes with frequency at different times of the year will be described.

The daily variations of phase and amplitude of the vertically incident wave on frequencies of 16 kc/s. (GBR) and 113 kc/s. and 70 kc/s. (Decca) are recorded at Cambridge, and the nature of the regular variations will be described by Mr. Bracewell.

The detailed knowledge available at vertical incidence is not yet matched by such detailed knowledge at oblique incidence. Mr. Tremellen will explain what is known from commercial experience on 16 kc/s., particularly with reference to atmospheric noise, and the possibilities of very low frequencies for navigational aids. Mr. Sanderson will discuss what is known about the phase variations of waves of frequency near 100 kc/s. observed at distances up to 300 miles by the use of Decca equipment, particularly with reference to the useful range of a navigation system on that frequency. Mr. Williams will discuss what can be deduced about the ionosphere from measurements made at different distances on Decca and other navigational aids working in the same frequency band.

In thinking of navigational aids we must investigate "phase diversity" effects. Published knowledge in this field will be summarized. Since phase diversity effects are produced by irregularities in the horizontal structure of the ionosphere there is an overlap with Saturday's session, and it is proposed that only observed effects should be discussed today, whereas their bearing on ionosphere theory should be discussed on Saturday.

Two attempts are being made at Cambridge to relate the behaviour at vertical incidence on 16 kc/s. to that at oblique incidence. In one the interference pattern formed between the ground wave and the ionospheric wave has been plotted in an aeroplane to a distance of 800 km. The results will be described in outline. Mr. Bracewell will describe the other experiment, in which detailed observations of amplitude observed at distances of 200 km. and 500 km. have been compared with those observed at a distance of 90 km.

In the discussion information derived from the observation of atmospherics and the signal strengths of long wave stations will be discussed.

Very Low Frequency Propagation

By S. B. SMITH and K. W. TREMELLEN *

Marconi's Wireless Telegraph Company Ltd.

The introduction of high-frequency communications during the years 1924/28, coupled with the poor operating economics of V.L.F. traffic, diverted industrial research to the 3–30 Mc/s. band. For many years academic research was not so extensive in the V.L.F. band, and only since 1944 has interest been renewed in this subject. This revival of interest was due to the failure of high-frequency ionospheric ray propagation to provide accurate radiolocation and to the belief that V.L.F. propagation may provide a more stable and accurate navigational means.

We propose to discuss some aspects of our earlier experiments during the years 1920/1926.

The contributions will be divided into the following sections :

(*a*) V.L.F. propagation, with particular reference to apparent abnormalities and other little known phenomena.

(*b*) World atmospheric noise centres, seasonal and diurnal variations in direction and intensity.

(*c*) Direction finding using different techniques, with special reference to reception along various geomagnetic paths.

* To be submitted to the Physical Society for publication.

The Effects of Sky-Wave on the Planning of Navigational Aids using Frequencies in the 70–130 kc/s. Band

By W. T. SANDERSON

Decca Navigator Company, Limited

The Decca system employs four stations which radiate continuously an unmodulated signal of about 1 kw. on frequencies 85·00, 127·5, 113·3 and 70·83 kc/s.—that is, in the ratio 6, 9, 8 and 5, which are phase-locked to a high degree of stability.

In the receiver the 3rd, 4th and 5th harmonics of the 85-kc/s. transmission are compared with the 2nd, 3rd and 6th of the other transmissions. The accuracy of the system depends on the stability of these patterns, and a number of observations has been made to determine the phase variation which can be expected at various ranges and times. Unlike the vertical incidence experiments, there is no absolute standard of phase against which to measure the variance of any transmission, but it is desirable to assess this factor so that the observed results at several points can be applied to find the expected pattern variance at any other point.

The results of some of the tests will be described. It appears that in English latitudes the night sky wave effects persist throughout the day during mid-winter on 70 kc/s., falling to about half the night value at 120 kc/s. During the summer daylight (April–September) the effects are so small up to 300 miles that they are difficult to measure, the standard error being less than a hundredth of a cycle.

Hawker has found that the secant of the sun's zenithal distance at noon gives a good approximation to the relative amplitude of the daylight errors throughout the year.

The implications of these results on the planning of medium range C.W. navigational aids will be outlined.

The Characteristics of Low-Frequency Radio Waves Reflected from the Ionosphere, with particular reference to Radio Aids to Navigation

By C. WILLIAMS

Royal Aircraft Establishment

The study of the performance of low-frequency radio navigation systems at great distances has provided useful data about the ionosphere. Using the Consol, Decca and POPI systems, records have been made of the changes in amplitude and phase which result from the reflection of the signals by the E layer of the ionosphere. The data are obtained at fixed monitor receiving points, and it is found that the phase variations which are produced by the continuously changing physical state of the ionosphere are a function jointly of the base-line distance between the transmitting stations and the distance of the receiver.

Other variations of phase depend on the difference between the ground- and ionospheric-wave paths and are most readily assessed from data collected with mobile receivers. Data collected with fixed receivers and with receivers in aircraft show the existence and the effect of both types of variations. From the airborne data it has been possible to establish the relative amplitude of the ground and ionospheric reflected wave as a function of distance from the transmitters and to estimate the height of the reflecting layer and the reflection coefficient at oblique incidence during day and night. Day values of 0·05 and night values of 0·23 have been obtained at frequencies near 100 kc/s.

Measurements on Long and Very Long Waves

By R. N. BRACEWELL
Cavendish Laboratory, Cambridge

This contribution describes the regular daily and seasonal variations of waves of 16 and 100 kc/s. observed after reflection from the ionosphere at distances of 90, 200 and 500 km.

A steeply incident wave from GBR Rugby (16 kc/s.), when received 90 km. away at Cambridge, shows that the reflection coefficient at night in all seasons is about 0·5. In winter this strong reflection persists throughout the day, but at sunrise in summer the reflection coefficient falls rapidly to a steady value of about 0·15. The equivalent height of reflection fluctuates about 90 km. at night in all seasons. From shortly before sunrise until noon the height h decreases. The total decrease is 18 km. in summer and 12 km. in winter, and except during twilight closely follows the expression $h=a \log \cos \chi+\text{const.}$, where χ is the sun's zenith distance and a is about 6 km. The daily decrease in height causes an average change of 500 degrees in the phase of the sky-wave. The polarization of the down-coming wave is left-handed and approximately circular throughout the day.

Observations of the less steeply incident reflections at a distance of 200 km. differ from the above only to the extent that in winter the twice-reflected wave becomes comparable with the once-reflected.

On the Decca frequencies (113 and 70 kc/s.) the reflection coefficient by day is of a markedly lower order of magnitude. During the night the reflection coefficient is 0·25 in winter and 0·15 in summer. About an hour after sunrise in summer the reflection coefficient begins to diminish and rapidly falls to 0·003, after which the signal is lost in the noise. In winter 1948 the reflection coefficient had the relatively high value of 0·1 during the day. In the two previous winters it was no more than 0·02.

The height of reflection begins to decrease at sunrise, and in the course of the three hours during which it can be measured falls about 7 km. The polarization remains constant over the sunrise period and is probably circular. The transition from summer to winter conditions occurred in November 1948 within the space of a week, but the return to summer conditions was much slower.

When GBR Rugby is observed at 500 km. distance, multiple reflections are important in determining the behaviour of the sky wave, and the reflection coefficient appears to be rather greater at this glancing incidence. There is a departure from the strict symmetry about noon which is observed at steeper incidence, especially in summer, when the afternoon recovery appears to lag. The sunrise change is more rapid, in fact its onset is often sharply defined within a minute or two. This change occurs well before sunrise at ground level and also well before the first changes which are noticed on steeper incidence. The time of its occurrence does not depend on the solar cycle, and can be predicted, for a given date, within about 10 minutes.

SESSION 2

THE REGULAR BEHAVIOUR OF MEDIUM AND SHORT WAVES

The Application of Ionospheric Data to Short-Wave Transmission Problems

By W. J. G. Beynon
University College, Swansea

A short survey is presented of the fundamental theory underlying the application of normal incidence ionospheric data to short-wave communication problems, with particular reference to calculating the maximum usable frequency (M.U.F.). Some aspects of the problem of applying normal incidence data on ionospheric absorption to the calculation of field strength in long distance transmission are also discussed.

A. F. Wilkins
Radio Research Station, Slough

No Title or Abstract received

(P', f) Records at Spitsbergen

By A. B. Whatman
Ministry of Supply

A selection of 60 (P', f) records will be shown which were made with Admiralty Type 249 equipment in Spitsbergen in 1942–43. These illustrate all the interesting normal and abnormal effects met with, and supplement those in a recent paper published in *Proc. Phys. Soc.* B, 1949, **62,** 307.

SESSION 3

THE IRREGULAR BEHAVIOUR ASSOCIATED WITH SOLAR EVENTS

Irregular Behaviour of the Ionosphere Associated with Solar Events

By W. R. PIGGOTT
Department of Scientific and Industrial Research

The phenomena occurring in the lower regions of the ionosphere which are associated with solar flares are briefly discussed. The types of experimental data obtained are summarized, and the geophysical significance of the results indicated.

The association of magnetic disturbances with solar flares and the main features of the corresponding ionospheric disturbance are described. It is shown that ionospheric and magnetic disturbance is not, in general, synchronous outside a zone of greatest storm activity. Two main types of ionospheric disturbance are identified:

(i) the quasi-auroral type of disturbance—type A;
(ii) the regular type of disturbance—type R.

A type A disturbance often includes severe disturbances in D and E regions (no reflection condition and auroral E), but its characteristic effects are in F region, where it produces an extremely rapid decrease in the ionization density, changes in height, followed by irregular variations and, if the layer is illuminated with light from the sun, a rapid recovery to normal or supernormal densities.

At night the A disturbance may be identified by the rapid recovery of the height towards normal values.

The fluctuations in the F region are closely associated with the fluctuations in the local magnetic field at the observing stations.

The type R disturbance consists of two phases: (i) a period during which the ionization density is above normal—the positive phase; (ii) a period during which the ionization density is below normal and shows a slow recovery superimposed on a regular diurnal variation—the negative phase. No connection between the magnetic variations and the R disturbance has been detected.

The relaxation time for type A is usually between a few minutes and one or two hours; for type R many hours to several days.

For a simple storm the disturbance varies fairly regularly with latitude in a given longitude zone, changing from type A with no type R in the active auroral zone, to type R with a little superposed type A, and then type R only. In the R zone the length of the positive phase increases and the depth of the negative decreases with increase of distance from the A zone until there may be no negative phase and, occasionally, no positive phase left. These disturbances are essentially unipolar and may reach to or beyond the equator. They are also most severe between a limited range of longitudes.

Some of the largest storms appear to consist of several disturbances, commencing at different times in different longitude or hemisphere zones.

The implications of these results are briefly considered.

Some Work at Cambridge on Radio Fade-outs

By K. WEEKES
Cavendish Laboratory, Cambridge

1. *Effects observed at vertical incidence on* 2·0 *and* 2·4 *Mc/s.*

Using a special technique, the changes of phase path and of amplitude of the pulse reflected from E region have been studied. Every fade-out is accompanied by simultaneous changes of the phase path which decreases 2–4 km. at the onset of the fade-out and increases again as the extra absorption decreases. The changes are different on the two frequencies used and the difference provides information about the distribution of ionization during the fade-out.

2. *Effects observed near vertical incidence on* 70 *and* 113 *kc/s.*

Using the C.W. transmissions of Decca, the sky wave has been isolated and its changes of phase and amplitude measured. The effects during a fade-out are very similar to those on higher frequencies. The phase path is reduced by an amount which may be as much as 20 km. and the amplitude decreases by a factor of at least 200 at the start of the fade-out. The phase and amplitude recover relatively rapidly as the intensity of the fade-out decreases.

3. *Effects observed near vertical incidence on* 16 *kc/s. and* 40 *kc/s.*

The outstanding fact in the study of very low frequencies is the large amplitude of the sky wave. During a fade-out the amplitude changes only slightly, and it is possible to study the phase changes throughout the fade-out. The phase path is reduced by an amount which may be as much as 20 km. and the amplitude changes by a factor of the order of 3. The phase changes appear to follow the solar flare intensity closely and to recover to the normal value rapidly when the fade-out is over. The effect on 40 kc/s. is similar to that on shorter waves.

4. An attempt will be made to outline an ionization distribution during the fade-out which would account for these observed results.

Les renforcements brusques des ondes très longues*

Par R. BUREAU

Laboratoire National de Radioélectricité

Read by R. RIVAULT, Université de Poitiers

Les renforcements brusques des ondes très longues ont été mis en évidence par l'enregistrement des atmosphériques; ils sont l'un des aspects des phénomènes radioélectriques consécutifs aux éruptions chromosphériques du soleil et sont ainsi d'un grand secours pour l'exploration indirecte de l'ionosphère. Grâce à eux, il a pu être dressé a posteriori une liste des perturbations ionosphériques à début brusque (P.I.D.B.) qui remonte à 1928, et qui met en évidence une influence très profonde du cycle solaire undécennal. Employé conjointement avec l'enregistrement du champ d'émetteurs sur ondes décamètriques, celui des atmosphériques met en lumière, dans la plupart des cas, le synchronisme des renforcements sur ondes très longues et des évanouissements brusques des ondes décamètriques. Sur les ondes supérieures à 16,000 mètres, les renforcements brusques s'amoindrissent ou disparaissent, ou se renversent même parfois.

Dans un certain nombre de cas, on peut associer une P.I.D.B. (due à un rayonnement solaire ultra-violet) au déclenchement ultérieur d'un orage magnétique (dû à un rayonnement corpusculaire) et à des troubles très profonds des radiocommunications; le délai qui s'écoule entre les deux phénomènes est en général voisin de 36 heures. Les P.I.D.B. associées aux observations d'éruptions chromosphériques et aux crochets du champ magnétique terrestre, deviennent un auxiliaire précieux dans les prévisions des orages ionosphériques.

Les études se poursuivent, de nouveaux domaines sont prospectés. Signalons des influences très rares, mais indubitables, du rayonnement cosmique; l'influence du seuil de sensibilité des enregistreurs. L'étude des anomalies de phases, entreprise à Cambridge sur les ondes très longues, complète heureusement la documentation relative aux P.I.D.B.

---

* This issue, p. 122.

SESSION 4

THE FORMATION OF IONIZED REGIONS

The Formation of the Ionized Regions

By K. WEEKES

Cavendish Laboratory, Cambridge

The variations of ionization density and height of maximum ionization for the E and F_1 regions agree well with the behaviour predicted by Chapman's simple theory of absorption of ultra-violet light in an isothermal atmosphere, provided it is assumed that the electrons are lost by a recombination process the rate of which is independent of the gas pressure. The value of the recombination coefficient deduced from the observations is surprisingly high compared with the theoretical estimates.

A closer comparison of the Chapman theory with the observations reveals small but important discrepancies, and the various assumptions underlying the theory are re-examined in an attempt to discover the cause of these discrepancies. In particular it is found that the assumption that the recombination rate is independent of the pressure is not essential if the layer is formed in a part of the atmosphere containing two gases which have different vertical distributions of partial pressure. Such a condition occurs near 100 km. when the oxygen is in transition from the molecular to the atomic state.

The gases which may be responsible for the production of the various layers are briefly discussed.

Theoretical Considerations Regarding the Formation of the Ionized Layers *

By D. R. Bates and M. J. Seaton

University College, London. (Read by D. R. Bates)

The detailed mechanisms involved in the formation of the E, F_1, F_2 and D layers by solar ultra-violet radiation are discussed. Use is made of the results of some recent calculations on the continuous absorption cross section of atomic oxygen and nitrogen, and of the evidence on the ionic composition of the layers that is provided by the analysis of the emission spectrum of the upper atmosphere during twilight. Certain difficulties are emphasized and attention is directed to measurements that would be helpful in elucidating the position.

* This issue, p. 129.

SESSION 5

IRREGULARITIES IN THE HORIZONTAL PLANE IN THE IONOSPHERE

Irregularities in the Horizontal Plane in Region E of the Ionosphere

By J. W. Findlay

Cavendish Laboratory, Cambridge

The fading of pulse signals returned from region E at vertical incidence occurs even when care is taken to receive only one of the two magneto-ionic components in the down-coming pulse. Experiments will be described by which the variations of the phase and amplitude of signals on a frequency of 2·4 Mc/s. have been measured. The rapidity of the fluctuations in phase of the signals can be assumed to be a measure of the irregularity of the reflecting region, and the variations of this irregularity with time of day and with season will be shown.

Experiments on the amplitude of pulse signals returned to three receiving points spaced about 100 metres apart on the ground have shown that the variations of amplitude may be explained by assuming that the irregularities in region E have at any time random motions among themselves and also an overall horizontal drift velocity. The theories which allow the velocities of these random motions and of the drift motion to be determined will be outlined. The results obtained from the observations at three receiving points will be interpreted by these theories, and the values derived from the observations for the r.m.s. velocity of the random motion and for the uniform drift velocity will be quoted.

Meteor Ionization in the Upper Atmosphere

By A. C. B. Lovell

Manchester University

Recent progress in the study of the scattering of radio waves from meteor trails makes it possible to assess within reasonable limits the total contribution of meteor ionization to the upper atmosphere and the irregularities thereby introduced. Significant changes from previous estimates, such as made by Skellett or Pierce, arise from two facts: (1) the discovery of the great daytime meteor radiants active from May to September; (2) the appreciation that only about 10^{-6} of the energy of the meteor is spent in ionization, and not the major part as previously assumed.

Observations on fairly high frequencies (e.g. 72 Mc/s.) show that the meteor trails are critically aspect sensitive at the instant of formation, but that the subsequent diffusion of the electron column gives rise to very complex types of radio wave reflections. On lower frequencies (e.g. 30 Mc/s.) a very much greater number of reflections is observed than would be anticipated for the scattering processes operative on the higher frequencies. These have been shown recently to be due to the loss of aspect sensitivity on the lower frequencies and the possible breaking up of the original column into separate clouds of ionization, thus producing multiple echoes instead of the single echo observed on the higher frequencies.

At least for frequencies above 25 Mc/s. it is thus established beyond doubt that all the transient echoes in the altitude range 80 to 120 km. are due to meteoric ionization, although there remains some doubt as to whether some of the fainter meteors may have an origin outside the solar system.

Scattering of Radio Waves from Region E

By G. Millington

Marconi's Wireless Telegraph Company

The existence of ionospheric irregularities was suggested over 20 years ago by T. L. Eckersley to explain the signals from high frequency transmitters received within the skip zone. For many years he carried out an extensive research into the scattering properties of the E layer.

The sporadic echoes scattered directly back from the E layer, unlike echoes observed at higher frequencies, show no definite correlation with visible meteors. These echoes, it is suggested, are only partly caused by meteors, and there must be another cause either in the sun or in the stars of the galaxy.

Long-distance scattering, in which the sources are illuminated, and the scattered signals received by waves reflected from the F layer, may be due to scattering centres in the E layer or on the ground. The experimental evidence must be critically examined to decide which is occurring.

Long-distance scattering can possibly cause radio reception over a route which is away from the great circle path when this is an impossible path owing to absorption or electron limitation. In the problem of frequency sharing in the H.F. band a scattering signal may be received from an unwanted station and cause interference with the wanted signal from another transmitter on the same frequency. The knowledge needed to assess likelihood of this type of interference will be discussed.

The Variations in Direction of Arrival of High-Frequency Radio Waves

By W. Ross

Radio Research Station, Slough

The variations in bearings observed on high-frequency radio transmissions which are discussed are those due to processes occurring in the propagation of the waves. The direction-finders which have been used are free from variable instrumental errors, such as polarization errors, which in early work tended to obscure the true bearing deviations.

High-frequency bearings are subject to continuous fluctuations. When rapid fluctuations are minimized by averaging over a few minutes, and when the "wide" bearings observed at a trough of a fading cycle are ignored, there remain slow fluctuations of periods from a few minutes up to about half an hour. These fluctuations have an amplitude comparable with the rapid fluctuations. They are similar in bearings taken over the same path on neighbouring frequencies and also in bearings taken at stations up to 10 km. apart.

These fluctuations are apparently due to a fairly localized tilting of the ionosphere at angles of a few degrees over distances of some tens of kilometres.

The rapid fluctuations, which are probably due to wave interference effects, are observed even when pulses are used, though they are smaller than with C.W. signals. This suggests that there are irregularities of the ionosphere of smaller scale than those responsible for the slow fluctuations.

At certain times, for example at sunrise, systematic ionospheric tilts are observed. It is suggested that the random irregularities may be due to the spreading of disturbances produced by perturbations in the regular diurnal changes in the ionosphere. (See *Nature, Lond.*, 1947, **159,** 132.)

Diffusion des échos su voisinage des fréquences critiques de F_2 *

Par R. Rivault

Université de Poitiers

On déduit de la théorie magnéto-ionique l'existence de trois composantes réfléchies par F_2, de fréquences critiques f_0, f_{X1} et f_{X2}. Les sondages à la verticale effectués à Poitiers comportent habituellement les composantes f_0 et f_{X1}, espacées de $f_H/2=0{,}6$ Mc/s. La composante f_{X2}, telle que $(f_{X1}-f_{X2})=f_H=1{,}2$ Mc/s., n'est enregistrée que très rarement. Mais l'on observe fréquemment des composantes de fréquences critiques comprises entre f_{X2} et f_{X1}, échelonnées suivant des valeurs sous multiples de f_H, 0,3 et 0,15 Mc/s. par exemple. Le plus souvent, le nombre de ces composantes est tel que l'intervale $f_{X1}-f_0$ est entièrement occupé par des réflexions diffuses. On peut admettre que c'est la même cause physique qui produit la multiplication de ces composantes et la diffusion.

L'analyse des enregistrements présentant des réflexions diffuses montre que ce phénomène ne se produit nettement qu'entre le coucher et le lever solaires, d'autant plus fréquemment et d'autant plus tôt dans la nuit que celle-ci est plus longue. Cette diffusion cesse au moment du lever du soleil au sol, un peu après le début de l'accroissement matinal des fréquences critiques, quand la stratification diurne reparait après le désordre thermique nocturne; elle n'est pas liée aux périodes magnétiquement perturbées ni à celles de grande activité solaire.

On rencontre une autre espèce de refléxions diffuses, enregistrées à des hauteurs virtuelles de 800 à 1200 kilomètres environ, sur des fréquences immédiatement supérieures aux fréquences critiques de F_2. Ces réflexions peuvent être intenses dans la deuxième partie de la nuit, et leur aspect net, principalement entre 3 et 5,2 Mc/s., fait penser à l'existence d'une région ionisée G supérieure à F_2. Elles cessent généralement après le lever du soleil, mais coincident souvent avec les périodes d'agitation magnétique et, surtout, avec les passages des principaux essaims de météorites.

* This issue, p. 126.

REVIEWS OF BOOKS

Fundamentals of Discharge Tube Circuits, by V. J. FRANCIS. Pp. x+133. 1st Edition. (London: Methuen & Co. Ltd., 1948). 6*s*. 6*d*.

There are very few books available on the fundamentals of discharge tube circuits. It is therefore very pleasing to the reviewer to see the present book which gives a survey of the whole subject.

No attempt is made to describe all the different circuits in common use, rather the author has attempted to discuss the more important fundamental properties of discharge tubes in relation to the circuits in which they are, or could be, operated. Chapters are included on the discharge tube as a circuit element, operation on D.C. supplies, dynamic characteristics, voltage and current wave-forms on A.C. supplies, operation on alternating current supplies, initiation and typical discharge lamp circuits and general principles of design.

The book should prove interesting and useful to a wide circle of readers particularly those who are actively concerned with the use and applications of discharge lamps. It is an authoritative account and can be recommended. D.T.

Optical Design and Lens Computation, by B. K. JOHNSON. Pp. viii+175. 1st Edition. (London: The Hatton Press Ltd., 1948.) 36*s*.

A former Lord Rayleigh once complained that every writer on Optics used a different notation and none of them read what the others had written. There is no doubt that this variety of notation has been the cause of the absence hitherto of any book dealing in a practical manner with the design of lens systems. Each professional lens designer has used his own notation and has been deterred from writing papers or books on the subject by the impossibility of using a universal language. Thus the old Optical Society died from want of matter to publish.

The late Professor Conrady, for the purpose of his lectures, employed a notation all his own, the attraction of which was the facility with which it could be reproduced on the typewriter and the duplicating machine, awkward as it was in other respects. His book, *Applied Optics and Optical Design*, introduced his notation and formulae to a wider circle than that of his students and a facsimile edition published in the U.S.A. during the late war carried it across the Atlantic. The Conrady notation and formulae are therefore known and can be easily referred to. Publication of his work was never completed ; his book is limited to telescope objectives and eyepieces. However, his lecture work continued on to photographic lenses and microscope objectives.

The present author makes use of the Conrady notation and deals in succession with the design of telescope objectives, eyepieces, photographic lenses and microscope objectives, reproducing the treatment Conrady gave in his lectures. The author pays generous tribute to Conrady's work.

By the design of optical systems should be implied a method of working out the forms of the constituents by anyone inexperienced in practice, that is, without a background of experience to start from. For this one needs an algebraic set of equations based on the aberrations, which may be deduced from the purely geometrical considerations of ray paths, by employing the principle of least time, or by consideration of the wave surface. The author uses none of these—his method is based on ray tracing. It is doubtful if the methods he displays are the best for training opticians, but they are presumably those taught in the Technical Optics Section of the Imperial College of Science and Technology, the only school of optical designing in the country.

The methods developed in this book will not appeal strongly to the mathematician, which is a pity, since there is a very great need to attract able mathematicians to the optical trade. The methods are directed rather to the student of limited mathematical attainments. The difficulty seems to be that mathematicians of originality are repelled by the drudgery of routine calculation. Those who are not repelled often lack originality.

The present book will have value therefore for students and may interest others in the subject and may even stimulate other designers to publish their techniques ; some of these may prove more attractive to mathematicians.

The book is well provided with examples : these are all worked out logarithmically, and are of course ray tracings. The clumsiness of design by ray tracing is well illustrated by the chapters on eyepieces and photographic lenses, where many pages are taken up by a treatment that could be dealt with algebraically in few. It is curious that the author deals exhaustively with the coma of eyepieces but is uninterested in field flatness, which is usually more important. Likewise, in his treatment of the photographic meniscus he is at pains to show how coma can be cured by means of the stop, while in practice the coma is used to cure the astigmatism. Some of the statements made are of doubtful validity and some of the suggestions of doubtful practical value : e.g. that flint glass should be tried for eyepiece lenses. It may be argued that this is an exercise for the student, but in a practical book it should be pointed out that this is not good practice, and why.

It is regrettable that the book is written in such bad English ; " alternately " for " alternatively " on p. 71 may be a printer's error, but such constructions as the following (chosen at random) abound : " Having determined the astigmatism . . . . the true curvature of field must now be determined. In order to do this, it is necessary . . . . " Unfortunately there is much scientific writing in this style.

The author is to be congratulated for his pioneering work in a difficult field. H. W. L.

Five-Figure Tables of Mathematical Functions, by J. B. Dale. 2nd Edition. Pp. viii + 121. (London : Edward Arnold & Co., 1949). Price 6*s*.

The first edition of this book must be familiar to most readers of this review. Since 1903 it has been one of the few convenient single-volume sets of miscellaneous tables compiled with the physicist and applied mathematician in mind. In addition to logs, antilogs and trigonometrical functions, it contained Bessel functions, gamma functions, elliptic integrals and Legendre functions, which are all of value at times to the ordinary physicist.

In the present edition the contents have hardly been altered; the table of logarithms is somewhat improved and there is now a table of the logarithms of hyperbolic functions. The old edition gave logs for the range 1·00(0·01)9·99, and the new gives, in addition, 1·000(0·001)2·999, which is a definite improvement. In the old edition the amount to be added for the next figure of the argument was given in the way that is usual with four-figure logarithms—columns for 1, 2, . . . 9, based on the mean differences of the line, were given opposite the entries. In the new edition this plan is retained for the first table, but for the second the actual differences are to be used, and there is a proper table of proportional parts, but, unfortunately, it is not on the same page as the logs themselves.

The whole book has been re-set, with modern " heads and tails " type, instead of the old, uniform-height type, and on paper which is very pleasant to handle. J. H. A.

Electronics in the Factory, edited by H. F. Trewman. Pp. 188. 1st Edition. (London: Pitman, 1949.) 20*s*.

This book is evidently intended to keep informed the works director or general manager—in short, the executive—as to the more general applications of electronics in industry. Examples of electronic devices for timing, counting, measurement, process control, eddy-current heating, etc., are discussed in simple terms which make a reasoned appeal to the imagination.

To the professional physicist the book makes light reading; I disposed of it in a restful evening and enjoyed it. G. I. F.

A Text Book of Heat, by LeRoy D. Weld. 1st Edition. Pp. x+436. (New York: The Macmillan Co., 1948). Price 25*s*.

This is described as a book for upper classmen, and has been based on the material used for 35 years as lecture notes in a junior-senior course in heat. This does not mean that it is old-fashioned, for modern work has been incorporated wherever it is needed. The properties of liquid helium receive due attention, the totally enclosed absorption type of refrigerator is described, and the International Temperature Scale properly treated.

The general outlook of the book is good. It aims evidently at presenting the student with principles rather than with practice, and many teachers may feel that it gives too little experimental detail; whether he is telling the student how to measure heat quantities or temperatures, or describing the evidence which led to the acceptance of the first law of thermodynamics, the author gives clear explanations of the principles of the experiment, but very little about the practical dodges adopted or required.

The scope of the book includes thermodynamics and statistical mechanics, but the latter suffers because quantum theory is always treated as something outside, to be superimposed on the classical theory. Apart from this, it is an excellent treatment.

The chapters are all accompanied with well selected sets of exercises, and with references for further reading. These show a distinctly American bias, which is, indeed, evident throughout the book. The attribution of the theory of light darts to Epstein may be an example of this, or may be a misprint.

As a whole, the book is very sound, and may give some very useful ideas and hints to those engaged in teaching this subject at the intermediate or pass degree level. J. H. A.

Theory of Dielectrics: Dielectric Constant and Dielectric Loss, by H. Fröhlich. Pp. vi+180. 1st Edition. (Oxford: Clarendon Press, 1949.) 18*s*.

The field of Professor Fröhlich's book is fairly precisely defined in its subtitle, "Dielectric Constant and Dielectric Loss": the contents being confined to theories of dielectric substances situated in ordinary (weak) fields for which there is no question of dielectric saturation or breakdown. The treatment, moreover, is entirely classical and there is no discussion of specifically quantal effects significant only at low temperatures (or for very strong fields). Within these limits, however, it deals with its subjects thoroughly.

There are four chapters. The first briefly defines the macroscopic concepts with which we are concerned; in particular, the ordinary static dielectric constant ϵ and the complex dielectric constant, $\epsilon_1+i\epsilon_2$, necessary in dealing with alternating fields. The quantities ϵ_1 and ϵ_2 are shown not to be independent functions of frequency; and the energy loss is proportional to ϵ_2.

Chapters II and III between them comprise over half the book, and are concerned with theories appropriate to the macroscopically static and dynamical cases, respectively. The former chapter deals with the Clausius-Mosotti formula, Onsager's extension of this, Kirkwood's theory of polar liquids and some rather general, and elegant, theorems due to the author. The applicability of these formulæ, to solids, liquids, gases, or solutions, is discussed in terms of the physical assumptions made: indeed, in the reviewer's opinion, one of the most useful features of the book is its classification of a wide range of dielectric problems in terms of the physical forces of primary importance. The latter chapter, concerned with relaxation effects, again classifies these according to their mechanisms, in particular distinguishing between loss due to elastically bound charges (resonance absorption) and that due to transitions of charges, or dipoles, over potential barriers. In the second case several models that lead to exponential decay functions are discussed in some detail, as also are the ensuing Debye equations. The effect on the frequency dependence of energy-loss of a spread of relaxation times is also investigated.

The final chapter deals with applications. In contrast to the previous chapters which had been heavy going, the reviewer found this rapid survey of a considerable range of experimental data very delightful reading. No attempt is made to include all existing data relevant to the theme of the book, but certainly the ground is well covered, and the chapter discusses such diverse topics as the dielectric constant of water (Kirkwood), the torsional flexibility of ketone molecules in paraffin solutions, and the ammonia absorption spectrum for centimetre waves (work of Bleaney and Penrose). This chapter, particularly, raises further problems for both theoretical and experimental work.

It ends, unfortunately, with a section, chiefly theoretical, on ionic crystals which is almost too condensed to be intelligible. Indeed the reviewer feels that perhaps the theoretical sections are, as a whole, rather over condensed and that sometimes greater clarity in the writing might have been achieved. It is possible that a more readable account would have resulted from a parallel development of the theory and the experimental material rather than a consecutive one. Despite this and a few minor defects (such as the omission of experimental points from Figure 19 and that the assumption that ϵ is independent of the field strength limits the general validity of some of the thermodynamical considerations) the book remains a very useful and stimulating addition to the literature.

The book is well produced, as one expects from the Oxford University Press ; and there is a useful list of nearly a hundred references to original papers. G. S. R.

CONTENTS FOR SECTION A

PAGE
Dr. H. Fröhlich and Mr. J. O'Dwyer. Time Dependence of Electronic Processes in Dielectrics . . . 81
Mr. J. H. Simpson. The Time Delay in Conduction and Breakdown Processes in Amorphous Solids . . . 86
Dr. E. Billig and Dr. P. T. Landsberg. Characteristics of Compound Barrier Layer Rectifiers . . . 101
Dr. J. V. Jelley and Dr. E. B. Paul. Fast Neutron Reactions with Fluorine and Sodium . . . 112
Mr. J. M. Hill and Mr. L. R. Shepherd. Slow Neutron-induced Activities in Europium and Samarium . . . 126
Mr. H. Elliot and Mr. D. W. N. Dolbear. Directional Measurements of the Diurnal Variation of Cosmic-Ray Intensity . . . 137
Dr. C. C. Butler, Mr. W. G. V. Rosser and Mr. K. H. Barker. Some Properties of Penetrating Cosmic-Ray Showers and Star Phenomena seen in the Cloud Chamber . . . 145
Miss B. Chowdhuri. Experiments on the Nature of Penetrating Events in Extensive Air Showers . . . 165

Letters to the Editor :

    Mr. F. K. Goward, Dr. E. W. Titterton and Mr. J. J. Wilkins. Observations of (γ, T), (γ, D) and (γ, np) Reactions in Boron . . . 172
    Mr. D. A. Wright and Dr. R. A. Weale. Conductivity of Thin Films . . . 173
    Dr. A. F. Gibson and Mr. T. S. Moss. The Photoconductivity of Bismuth Sulphide and Bismuth Telluride . . . 176
    Mr. J. J. Wilkins. Validity of Two-dimensional Design of Synchrotron Pole-Faces . . . 177
    Mr. R. Bowers and Dr. K. Mendelssohn. Properties of Superflow in Liquid Helium II . . . 178
Reviews of Books . . . 181
Contents for Section B . . . 182
Abstracts for Section B . . . 183

ABSTRACTS FOR SECTION A

Time Dependence of Electronic Processes in Dielectrics, by H. FRÖHLICH and J. O'DWYER.

ABSTRACT. In solid dielectrics electrons may be excited from low levels into conduction levels or into traps with energies slightly lower (shallow traps) by an ionizing radiation or by the action of strong fields. It is assumed that collisions amongst electrons in these levels are so frequent that an electronic temperature T is established which may be different from the lattice temperatures. Also the number of electrons in these levels need not correspond to equilibrium at temperature T, in which case transitions to or from lower levels (of lattice defects or foreign admixtures) will take place. Such transitions are either due to excitation by sufficiently fast-moving electrons in the conduction levels, or to the inverse process. In this way it is possible to account for recombination within a reasonable time provided these lower levels (denoted as deep traps) are not much more than 1 electron volt below the lowest of the shallow traps. The difficulty of having to introduce recombination with simultaneous emission of very many vibrational quanta, an extremely unlikely process, is thus avoided. Other applications of this method, for example to the time dependence of electronic conductivity in strong fields, are indicated.

The Time Delay in Conduction and Breakdown Processes in Amorphous Solids, by J. H. SIMPSON.

ABSTRACT. A method of estimating the time delay in changes of conductivity and in the breakdown process for amorphous materials, based on H. Fröhlich's high-temperature breakdown theory, is given. Numerical values of these times are obtained for the case of soda-lime glass, and it is estimated that the time constant of the conduction process is between 10^{-4} and 10^{-5} second, while that of the breakdown process may be considerably shorter. Experimental data obtained by Turner and Lewis indicate that the estimate for the conductivity process is of the correct order.

Characteristics of Compound Barrier Layer Rectifiers, by E. BILLIG and P. T. LANDSBERG.

ABSTRACT. The assumptions involved in Mott's and Schottky's theories of rectification are analysed and a potential barrier at the metal–semiconductor interface is proposed which enables one to pass continuously from one theory to the other. The model is directly applicable to the practically important case of rectifiers possessing an additional insulating layer between semiconductor and counter-electrode. The current–voltage characteristic of such a compound barrier is obtained and contains Mott's and Schottky's results as special cases. In this connection, an expression for the volume distribution of current carriers is obtained and plotted for the case of a linear potential distribution in the barrier. A discussion of the distribution of charges throughout the rectifier leads naturally to the concept of a capacitance of the barrier, and to the derivation of the capacitance–voltage characteristic. In this connection, a definite experimental problem is suggested by the analysis.

Fast Neutron Reactions with Fluorine and Sodium, by J. V. JELLEY and E. B. PAUL.

ABSTRACT. The yields of the (np) and (nα) reactions in fluorine and sodium have been measured by observations on the radioactivity produced by bombardment with neutrons from the (dn) reactions in Li, B, Be and D. Cross sections have been estimated and methods of deriving the absolute disintegration rates and the neutron flux are described.

The reaction energies have also been derived, by assuming a particular form for the excitation function, and using published data on the energy distribution of the neutrons used.

Slow Neutron-induced Activities in Europium and Samarium, by J. M. HILL and L. R. SHEPHERD.

ABSTRACT. The activities produced by the neutron irradiation of europium and samarium have been investigated by means of a thin lens type β-ray spectrometer and by absorption and coincidence methods. The 9·2-hour activity in europium is due to the dual decay of ^{152}Eu by β^- emission and by K capture. The long period activity is due to both ^{152}Eu and ^{154}Eu, and again both β^- emission and K capture are present. This long period K capture is due to ^{152}Eu, and the β^- emission is possibly due to both ^{152}Eu and ^{154}Eu. The energies of the emitted radiations are discussed and possible decay schemes proposed.

The 47-hour activity induced in samarium is due to the β^- decay of ^{153} Sm, and measurement of the energies of the emitted radiations has been made.

Directional Measurements of the Diurnal Variation of Cosmic-Ray Intensity, by H. ELLIOT and D. W. N. DOLBEAR.

ABSTRACT. Measurements of the solar daily variation in cosmic-ray intensity have been made for a total of 360 days using two arrays of Geiger–Müller counters inclined at 45° to the vertical in the North and South directions. Both the diurnal and semi-diurnal variations are found to be significantly different for the two directions, and since the particles recorded by the two sets have passed through the same amount of atmosphere under similar conditions, it must be supposed that at least a part of the variation is due to causes which lie outside the earth's atmosphere.

It is found that both the amplitude and phase of the daily variation for the South-pointing recorder vary with season, whereas in the North direction the amplitude only shows any seasonal variation. The amplitude is greatest in summer in both directions.

The experimental results are compared with those of earlier workers and are discussed in relation to some of the interpretations of the daily variation which have been put forward. The difference between the variations in the two directions is interpreted as evidence for a non-isotropic distribution of the primary cosmic rays in space. A variation with sidereal and not with solar time would be expected under these conditions, but it is pointed out that the magnetic field of the sun may modify the direction of arrival of the primary rays at the earth so as to produce a resultant solar time variation.

Some Properties of Penetrating Cosmic-Ray Showers and Star Phenomena seen in the Cloud Chamber, by C. C. BUTLER, W. G. V. ROSSER and K. H. BARKER.

ABSTRACT. Cloud-chamber photographs of cosmic-ray showers containing penetrating particles have been obtained at sea level. The spectrum of the penetrating particles has been measured and a study made of the interaction of the energetic particles in lead and in the gas of the cloud chamber. Conclusions about the nature of the penetrating particles are discussed; these show that about half the penetrating particles are protons and the remainder are probably π-mesons and mesons of heavier mass. Three heavy particles produced in stars are probably τ-mesons.

Experiments on the Nature of Penetrating Events in Extensive Air Showers, by B. CHOWDHURI.

ABSTRACT. The discharges, in coincidence with extensive showers, of individual counters of a group placed in a cavity under 15 cm. lead have been studied with particular reference to the number of these counters simultaneously discharged in a single event. It is shown that the observed multiplicity of discharge is not consistent with operation by particles, each discharging one counter only, distributed at random in the extensive shower falling on the apparatus. The multiplicity is also inconsistent with that expected for low energy photons at the extreme range of normal cascades, and with that arising from the secondary electrons of a reasonable μ-meson spectrum. Entry of energy into the shielded cavity is shown to take place over an area much smaller than the total area of the cavity, but the mechanism then leading to the observed multiplicity has not been identified.